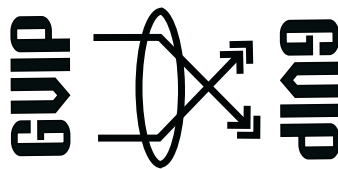


The International Congress for global Science and Technology



ICGST International Journal on Graphics, Vision and Image Processing (GVIP)

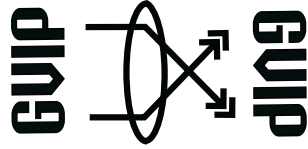
**Volume (12), Issue (I)
April 2012**

**www.icgst.com
www.icgst-amc.com
www.icgst-ees.com**

GVIP Journal
ISSN Print 1687-398X
ISSN Online 1687-3998
ISSN CD-ROM 1687-4005
© ICGST LLC, Delaware, USA, 2012

Table of Contents

Papers	Pages
P1151142873, R. Harrabi and E. Ben Braiek, “Color Image Segmentation based on the Optimal Multilevel Thresholding Technique”	1--8
P1151151088, Anh Tu Tran, Koichi Harada, “Multi-view Video Segmentation based on Bayesian Estimation and Graph Cut”	9--14
P1151205107, Nahed Solouma, Ph.D. and Amro El-Dib, MD, “Accurate Detection of Microaneurisms and Hemorrhage for the Planning of Laser Treatment”	15--21
P1151209132, Hassène Seddik and Ezzeddine Ben Braiek, “Optimal Watermarking Scheme in Time Frequency Domain for most Robustness and better Imperceptibility”	24--29
P1151215168, L. Sadouki , B. Haddad, D. Kemikem, “Modeling of the precipitations in radar and satellite images using the Autoregressive process”	31--36
P1151010018, Hassen Lazrag, Med Ali Hamdi and Med Saber Naceur, “2D Segmentation of Intravascular Ultrasound Images: A Fast-Marching Method”	37--41
P1151201093, Bulusu Rama and Jibitesh Mishra “Using three-dimensional (3D) Sierpinski Gasket to Generate and Recursively Re-Generate 3D Fractals - Closing the Self-Similarity Loop”	43--48
P1151150084, H. Essid, A. Sellami, V. Barra, I.R. Farah, “Monitoring intra-urban changes with Hidden Markov Models using the spatial relationships”	49--55



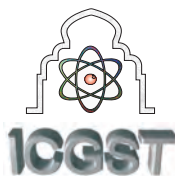
**ICGST International Journal on Graphics, Vision and Image
Processing
(GVIP)**

**A publication of the International Congress for global
Science and Technology - (ICGST)**

ICGST Editor in Chief: Dr. rer. nat. Ashraf Aboshosha

www.icgst.com, www.icgst-amc.com, www.icgst-ees.com

editor@icgst.com



Color Image Segmentation based on the Optimal Multilevel Thresholding Technique

R. Harrabi and E. Ben Braiek

Univ. Tunisia, CEREP Unit Research, ESSTT 5 Av. Taha Hussein 1008 Tunis, Tunisia

refka_esstt@yahoo.fr, ezzedine.benbraiek@esstt.rnu.tn

Abstract

Automatic thresholding has been widely used in the domain of image processing for automated segmentation images. The Otsu's method is a commonly used thresholding technique. It provides satisfactory results for thresholding images with bimodal distribution histogram, but it is impractical when extended to multilevel thresholding. In this paper, a new automatic thresholding method for colour image segmentation called the TSMO (Two-Stage Multi-level Thresholding) is presented. We revised the Otsu method for selecting optimal threshold values for both unimodal and bimodal distributions, and tested the performance of the revised method, on the colour images segmentation. This algorithm is iterative and outperforms Otsu's method by greatly reducing the iterations required for computing the between-class variance in an image. The experimental results for synthetic and biomedical colour images demonstrate the success of the proposed method, compared to many other methods.

Keywords: Thresholding, Otsu's method, two stage multi-thresholding, colour image segmentation, space colour.

1. Introduction

The problem of image segmentation has received considerable attention in the literature [1], [2], [3]. It is one of the most important tasks which determine the quality of the final result of image analysis. In this framework, colour image segmentation has wide applications in many areas [4] [5], and several methodologies have been proposed to tackle this problem [6] [7] [8].

In their recent work, Cheng et al. [9] show that the color image segmentation can be considered as a pixel labeling process in the sense that all pixels belonging to the same homogeneous region are assigned to the same label while according to similar color. In color image segmentation, the pixel color is given as three values corresponding to the three component images R (Red), G (Green) and B (Blue). Many papers dealing with segmentation using different kinds of color spaces have been developed by

several authors [9] [10], such as RGB space or their transformations (linear/non linear). Each color representation has its advantages and disadvantages.

Nonlinear color transformations such as HSI and normalized color space have essential singularities which are non-removable, and there are spurious modes in the distribution of values resulting from nonlinear transformations. RGB is suitable for color display, but not good for color scene segmentation and analysis because of the high correlation among the R, G, and B components [11] [12].

Hence, linear spaces are very difficult to discriminate highlights, shadows and shading in color images. Using HSI can solve this problem to some extent except that hue is unstable at low saturation [13]. In fact, the main problem of the color image segmentation is the choice of the adapted color model for a specific application.

Thresholding is widely used in many image processing applications such as (1) medical image applications [14]; (2) automatic visual inspection of defects [15] [16]; (3) optical character recognition [17], and (4) detection of video change [18] [19]. Thresholding is the simplest technique which involves the basic assumption that the objects and backgrounds in an image have distinct gray-level distributions. Many global thresholding techniques have been developed over the past years to segment images and recognize patterns [20] [21].

Otsu's method, as proposed in [20], is based on discriminant analysis. The threshold value is selected by maximizing the separability of the classes in gray levels. This method provides satisfactory results when thresholding an image with a histogram of bimodal distribution, but it presents several challenging difficulties for multilevel thresholding selection in images. Otsu's method can be easily extended to a multilevel threshold problem, but it is inefficient in determining the optimal thresholds, especially if there is a large class number M required in an image. To improve the efficiency of Otsu's method, many methods have been proposed. For example, Liao et al. [22] presented a modified version of Otsu's method called "recursive Otsu method" to find the threshold values by accessing



the pre-computed modified between-class variance through a look-up table (i.e., H-table). Moreover, Fan and Lin [23] proposed a hybrid optimal estimation algorithm to deal with the multi-level thresholding problem. With the same objective, Quweider et al. [24] presented a new algorithm based on dynamic programming and the optimal partitioning of the image data space on an interval of gray levels, which is commonly used in single and multi-level thresholding problems. This method can reduce the number of gray levels from a fine to coarse fashion, and is shown to offer very good results when compared to many existing methods.

To significantly improve the deficiencies in Otsu's method with regard to selecting the multi-level threshold, Deng-Yuan Huang et al. [25] proposed an algorithm called the Two-Stage Multithreshold Otsu method (TSMO). The TSMO method outperforms Otsu's method by greatly reducing the iterations required for computing the between-class variance in an image.

In this paper an investigation of the automatic thresholding technique is described. We use the TSMO method for segmenting the color images. So, this work may be seen as an extended of the method proposed by Deng-Yuan Huang et al. [25] for the color images. In their paper, the authors presented a fast algorithm of the gray level image segmentation, which is a modified version of the conventional Otsu method, but it determines the multi-level threshold in a two-stage fashion. Hence, this paper is devoted to the determination of the multi-level threshold, applied to color images segmentation that contains more than two classes. The idea is based on the histogram thresholding of the Hue component in the HSI color space. The concept of the Two-stage multithreshold Otsu's method (TSMO method) was discussed in [25], which is used to finding the optimal multi-level threshold in a two-stage fashion.

The rest of this paper is structured as follows: Section (2) briefly describes the conventional Otsu's method. Section (3) introduces the proposed method for color images segmentation. The experimental results are discussed in section (4), and the conclusion is given in section (5).

2. Conventional Otsu method

The Otsu's method [20], is good for thresholding a histogram with bimodal or multimodal distribution. This method, however, fails if the histogram is unimodal or close to unimodal.

Assume that an image can be represented in L gray levels ($1, 2, \dots, L$). The number of pixels at level i is denoted by f_i and N represents the total number of pixels in a given image. For a given gray-level image, the occurrence probability of gray level i is defined as:

$$p_i = \frac{f_i}{N}; \quad p_i \geq 0; \quad \sum_{i=1}^L p_i = 1 \quad (1)$$

In the case of single thresholding, the pixels of an image are divided into two classes, C_1 and C_2 , by a threshold at level t , where class C_1 consists of gray levels from 0 to t , and class C_2 contains the other gray levels with $t+1$ to L .

The cumulative probabilities (w_1 and w_2) of the two classes, respectively, are given by:

$$w_1 = \sum_{i=1}^t p_i \quad (2)$$

$$w_2 = \sum_{i=t+1}^L p_i \quad (3)$$

The mean levels (μ_1 and μ_2) for classes C_1 and C_2 , can be computed as:

$$\mu_1 = \sum_{i=1}^t \frac{ip_i}{w_1} \quad (4)$$

$$\mu_2 = \sum_{i=t+1}^L \frac{ip_i}{w_2} \quad (5)$$

Otsu [20] selects an optimal threshold t^* that maximizes the between-class variance σ_B^2 in equation (6) where μ_T is the mean intensity of the original whole image.

$$t^* = \arg \max_{1 \leq t \leq L} \left\{ \sigma_B^2(t) \mid \sigma_B^2 = w_1(\mu_1 - \mu_T)^2 + w_2(\mu_2 - \mu_T)^2 \right\} \quad (6)$$

and

$$\mu_T = \sum_{i=1}^L ip_i \quad (7)$$

The Otsu's method [20], described here can be easily extended to image multilevel thresholding [22].

3. Proposed method

The Otsu's method [20] still remains one of the most referenced thresholding methods. This method selects threshold values that maximize the between-class variances of the histogram. It gives satisfactory results when the numbers of pixels in each class are close to each other. However, with an increase of the number of classes in an image, this method becomes rather inefficient because it requires a large number of iterations to compute the cumulative probability (zeroth-order moment) and the mean (first-order moment) of a class.

Hence, this paper is devoted to selecting the multi-level threshold, applied to color image segmentation, where, we aim at providing a help to the doctor for the follow-up of the diseases of the breast cancer. The objective is to rebuild each cell from the Hue component of the original image represented in the HSI color space. So, this work may be seen to be straightforwardly complementary to that in the paper proposed by Deng-Yuan Huang et al. [25], and the generalization of this method to color images segmentation.

HSI system separates color information of an image from its intensity information. Color information is represented by hue and saturation values, while intensity, which describes the brightness of an image, is determined by the amount of the light. Hue is the most useful attribute in color segmentation since it is less influenced by the nonuniform illumination such as haze, shadow, or reflect lights [26]. Hue can be obtained by a nonlinear transformation from Red, Green and Blue color features [26]:



$$\text{Hue} = \arctan\left(\frac{\sqrt{3}(G - B)}{(R - G) + (R - B)}\right) \quad (8)$$

Hue represents basic colors, and is determined by the dominant wavelength in the spectral distribution of light wavelengths. It reflects the predominant color of an object and has a great capability in subjective color perception [26].

To significantly improve the deficiencies in Otsu's method with regard to selecting the multi-level threshold, we use an algorithm called the Two-stage multithreshold Otsu's method (TSMO method).

A general concept of the TSMO method is given in Ref. [25]. The idea of this method is quite simple and straightforward: to greatly reduce the iterations required for calculating the zeroth and first-order moments of a class. In the first stage of the TSMO method, the histogram of an image with L gray levels is divided into M_z groups which contain N_z gray levels.

Let Ω denote the groups of the total image space; then $\Omega = \{\Omega_j | j = 0, 1, \dots, M_z - 1\}$, where j represents the group number. Hence, each group contains a certain range of gray levels as follows: Ω_0 contains a range of gray levels $\{0, 1, \dots, N_z - 1\}$, Ω_1 with gray levels $\{N_z, N_z + 1, \dots, 2N_z - 1\}$, ..., Ω_q with gray levels $\{qN_z, qN_z + 1, \dots, (q+1)N_z - 1\}$, ..., and the last group Ω_{M_z-1} with gray levels $\{(M_z - 1)N_z, (M_z - 1)N_z + 1, \dots, M_z N_z - 1\}$.

The number of cumulative pixels (the zeroth-order cumulative moment), in the q^{th} group denoted by g_{Ω_q} can be calculated as:

$$g_{\Omega_q} = \sum_{i=q \times N_z}^{(q+1) \times N_z - 1} f_i \quad (9)$$

where f_i represents the number of pixels with gray level i . Since each group contains N_z gray levels, the corresponding gray level value for each group can be considered as a mean value for those N_z gray levels. Therefore, the corresponding gray level value or mean intensity (the first-order cumulative moment), in the q^{th} group denoted by i_{Ω_q} can be calculated as:

$$i_{\Omega_q} = \frac{\sum_{i=q \times N_z}^{(q+1) \times N_z - 1} i f_i}{\sum_{i=q \times N_z}^{(q+1) \times N_z - 1} f_i} \quad (10)$$

$$= \frac{1}{g_{\Omega_q}} \sum_{i=q \times N_z}^{(q+1) \times N_z - 1} i f_i$$

Hence, Otsu's method can be applied to find the optimal threshold j^* by maximizing the between-class variance (σ_B^2) with the sets of i_{Ω} and g_{Ω} . The optimal threshold j^* which is also regarded as the number of the group into which the maximum variance of the between-class, i.e., $(\sigma_B^2)_{\max}$, falls with the corresponding group Ω_{j^*} is defined as:

$$j^* = \arg \max_{0 \leq j \leq M_z - 1} \{(\sigma_B^2(j))\} \quad (11)$$

If an image can be divided into two classes, C_1 and C_2 , by Ω_{j^*} , where class C_1 consists the group from Ω_0 to Ω_{j^*} , and class C_2 contains the other groups with Ω_{j^*+1} to Ω_{M_z-1} , then the numbers of the cumulative pixels and the means for the two classes, respectively, are given by:

$$w_1(j^*) = \sum_{j=0}^{j^*} g_{\Omega_j} \quad (12)$$

$$w_2(j^*) = \sum_{j=j^*+1}^{M_z-1} g_{\Omega_j} \quad (13)$$

and

$$\mu_1 = \frac{S_1}{w_1} \quad (14)$$

$$\mu_2 = \frac{S_2}{w_2} \quad (15)$$

where S_1 and S_2 are the first-order cumulative moments for classes C_1 and C_2 , respectively

$$S_1 = \sum_{j=0}^{j^*} i_{\Omega_j} g_{\Omega_j} \quad (16)$$

$$S_2 = \sum_{j=j^*+1}^{M_z-1} i_{\Omega_j} g_{\Omega_j} \quad (17)$$

Thus, the maximum variance of the between-class $(\sigma_B')^2_{\max}$, can be easily found using the modified version of Otsu's method proposed by Liao et al. [22]. In the case of bi-level thresholding ($M = 2$), the maximum variance of the between-class is defined as:

$$\begin{aligned} (\sigma_B')^2_{\max} &= \sum_{k=1}^M w_k \mu_k^2 = w_1 \mu_1^2 + w_2 \mu_2^2 \\ &= \frac{S_1^2}{w_1} + \frac{S_2^2}{w_2} \\ &= \frac{S_1^2}{w_1} + \frac{(S_T - S_1)^2}{N - w_1} \end{aligned} \quad (18)$$

where M is the class number in an image, S_T is the sum of S_1 and S_2 , and N is the total number of pixels in an image. That is

$$\begin{aligned} S_T &= S_1 + S_2 \\ N &= w_1 + w_2 \end{aligned} \quad (19)$$

In the second stage of the TSMO method, since Ω_{j^*}

contains the gray levels with $(j^*)N_z$ to $(j^* + 1)N_z - 1$ in which $(\sigma_B')^2_{\max}(j^*)$ occurs has already been found in the first stage, Otsu's method can be applied again to group Ω_{j^*} in a similar fashion to found the optimal threshold t^* . Hence

$$t^* = \arg \max_{(j^*)N_z \leq t \leq (j^* + 1)N_z - 1} \{(\sigma_B')^2(t)\} \quad (20)$$



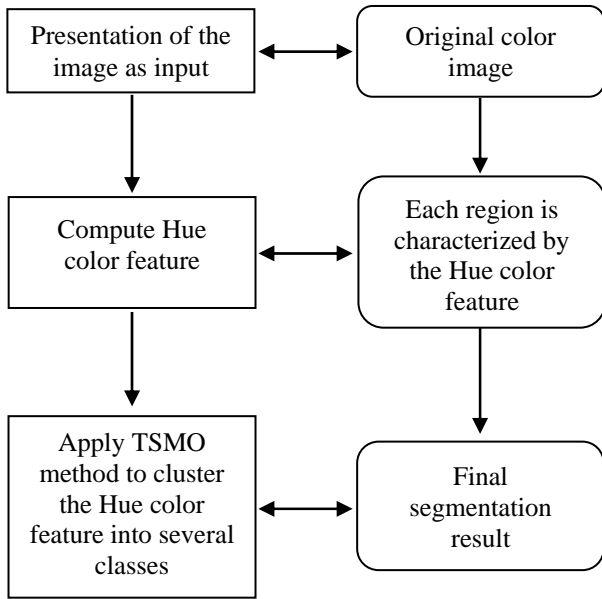


Figure 1. Flowchart of the proposed method

In fact, the first phase of the proposed approach is the determination of the color feature hue. In the second phase, the TSMO method is used to cluster the feature Hue into several classes with every class corresponding to one region in the segmented image.

Therefore, the pixels are divided into several groups with each group having similar colors. The proposed method can be described by a flowchart given in Figure 1.

4. Experimental results

The described method has been extensively tested on color cells images, where, we aim at providing a help to the doctor for the breast cancer follow-up diseases. Some experimental results are shown in Figures 4-7. The images are originally stored in RGB format. Each of the primitive colors (red, green and blue) is represented by 8 bits and has an intensity range from 0 to 255. We applied the proposed method to both RGB and HSI color spaces and compared the results with those obtained by using the Otsu's method and Fuzzy approaches.

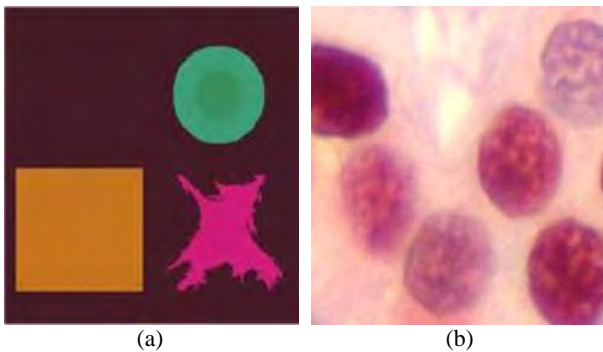


Figure 2. Synthetic image (a) and real medical cells image (b)

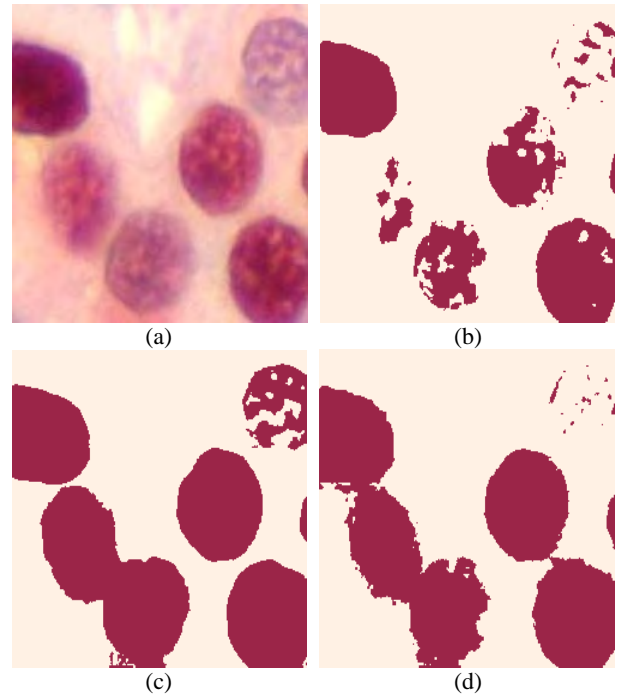


Figure 3. Segmentation results on a color image, (a) Original image (256x256x3) with gray level spread on the range [0,255]. (b) Red resulting image by TSMO-32 method. (c) Green resulting image by TSMO-32 method. (d) Blue resulting image by TSMO-32 method. The selected thresholds are 201, 116, and 122 respectively.

The two images shown in Figure 2 were used in order to illustrate the stages of the segmentation algorithm and visually assess the quality of the segmentation results. The synthetic image (Figure 2(a)) contains 5 areas and can be considered as piecewise constant in most of its areas, the background intensity level is 70, 28 and 39 for the red, green and blue components, respectively. Figure 2(b) shows a real medical cells image, where the noise statistically tested and found to be approximately additive Gaussian.

Figure 3 shows the segmentation result based on the described algorithm of the image Figure 2(b) represented in the RGB color space applied to the red, green and blue color features, respectively. In this case, a region is recognized in green component but is not identified by red and blue components. This shows that the RGB color space has a strong correlation of its three components, and hence, the use of a single information source leads to bad results.

Comparing the results, we can find that the cells are much better segmented in Figure 3(c) than those in Figure 3(b) and Figure 3(d). Also, the first resulting image contains some holes in the cells, which do not exist in another resulting image. This demonstrates the necessity to choose the color model and the segmentation algorithm adapted for a specific application.

For purpose of comparison, we apply the proposed approach and some Fuzzy and automatic thresholding approaches to the same color image segmentation, using HSI color space. The latter methods include those of N. Otsu [20], Liao et al. [22], Lim et al. [27] and Liew et al. [28].

The algorithm proposed by Lim et al. [27] called (TFCM), is based on the Thresholding and the Fuzzy C-means (FCM) techniques.



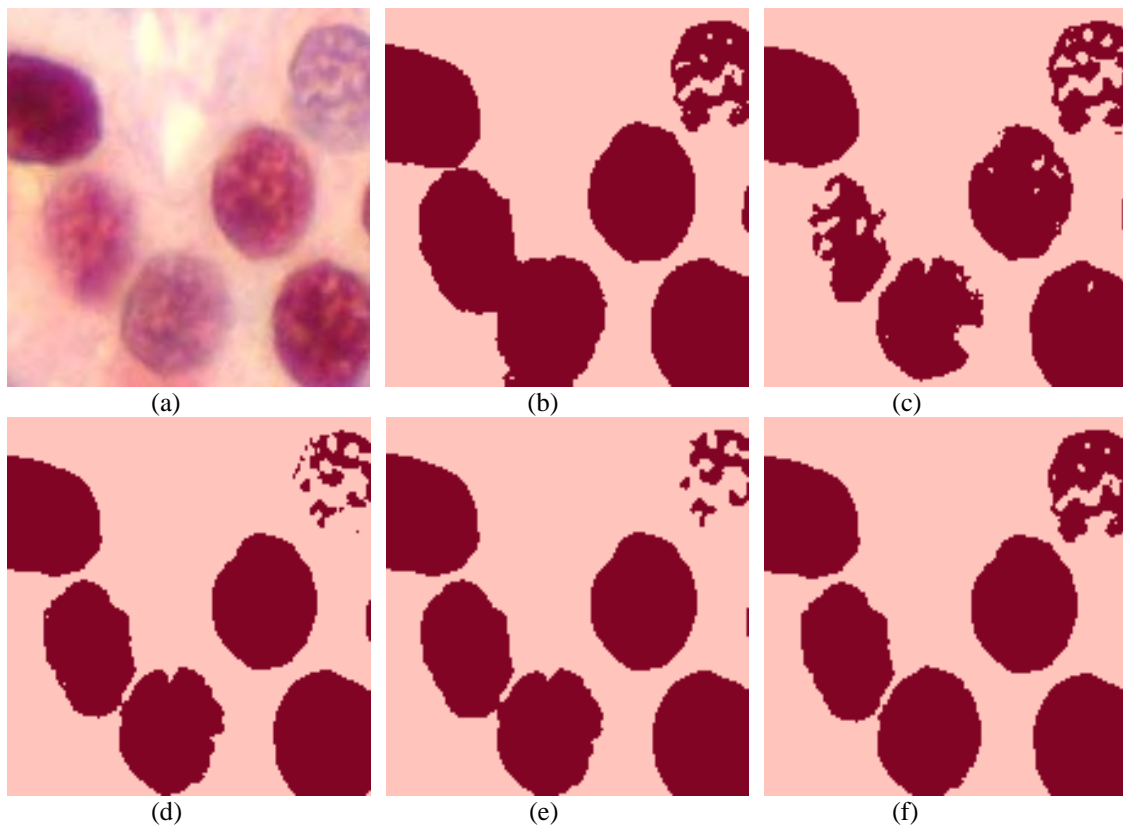


Figure 4. Comparison of the proposed segmentation method with other existing methods on a medical image (2 classes, various cells), (a) original image ($256 \times 256 \times 3$): color medical image with RGB description, (b) Hue resulting image by the HCM algorithm, (c) Hue resulting image by the FCM algorithm, (d) Hue resulting image by the TFCM algorithm, (e) Hue resulting image by the MFCM algorithm, (f) Hue resulting image by the proposed TSMO method.

Table 1. Segmentation sensitivity From HCM, FCM, TFCM, MFCM and TSMO Methods for the Data set Shown in Figure 5

	HCM	FCM	TFCM	MFCM	TSMO
<i>SENSITIVITY SEGMENTATION (%)</i>					
<i>Image 1</i>	63.61	74.75	76.82	78.45	87.96
<i>Image 2</i>	62.26	73.16	75.18	76.79	86.09
<i>Image 3</i>	65.23	74.30	76.34	77.98	87.43
<i>Image 4</i>	63.08	74.12	76.17	77.79	87.22
<i>Image 5</i>	63.26	74.33	76.39	78.02	87.47
<i>Image 6</i>	83.67	87.99	89.23	89.35	89.38
<i>Image 7</i>	66.50	73.44	75.47	77.08	86.42
<i>Image 8</i>	69.11	74.15	76.20	77.83	87.26
<i>Image 9</i>	67.22	78.99	81.17	82.90	92.95
<i>Image 10</i>	71.73	84.28	86.62	88.46	99.18
<i>Image 11</i>	72.12	84.74	87.08	88.94	99.71
<i>Image 12</i>	71.66	82.97	91.83	93.78	95.32

The methodology uses a coarse-fine concept to reduce the computational burden required for the FCM algorithm. Furthermore, to greatly improve the efficiency of FCM method, Liew et al. [28] proposed a new algorithm called the MFCM method. In fact, the MFCM method outperforms FCM method by using the local spatial information. This method improve the classification results and provides satisfactory segmentation results.

However, the segmentation results obtained by the conventional Otsu method, the fuzzy algorithms and the proposed TSMO method as described earlier, are shown in Figures 4 and 5.

Figure 4 shows the segmentation results obtained by the HCM algorithm [29], the FCM algorithm [5], the TFCM method [27], the MFCM method [28] and the proposed

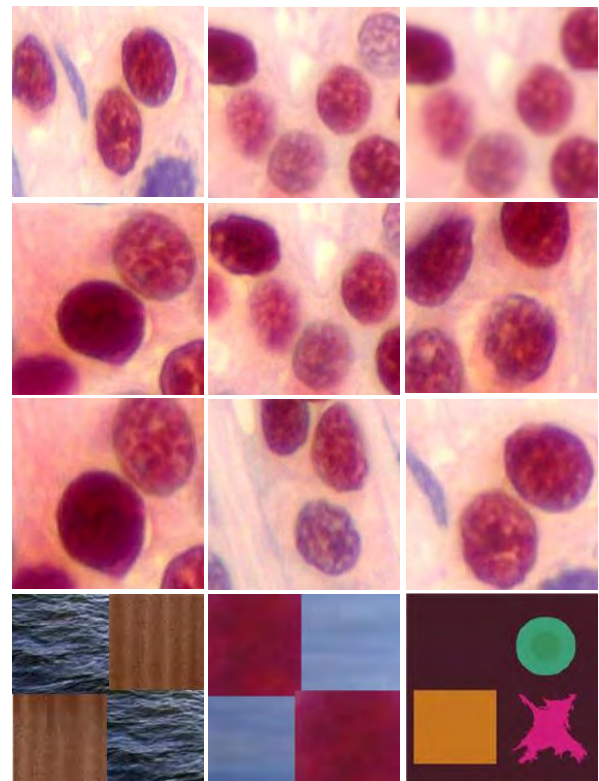


Figure 5. Data set used in the experiment. Twelve images were selected for a comparison study. The patterns are numbered from 1 to 12, starting at the upper left-hand corner.

TSMO method applied to the color feature Hue of the Figure 2(b). They correspond, respectively, to Figure 4(b), Figure 4(c), Figure 4(d), Figure 4(e) and Figure 4(f).



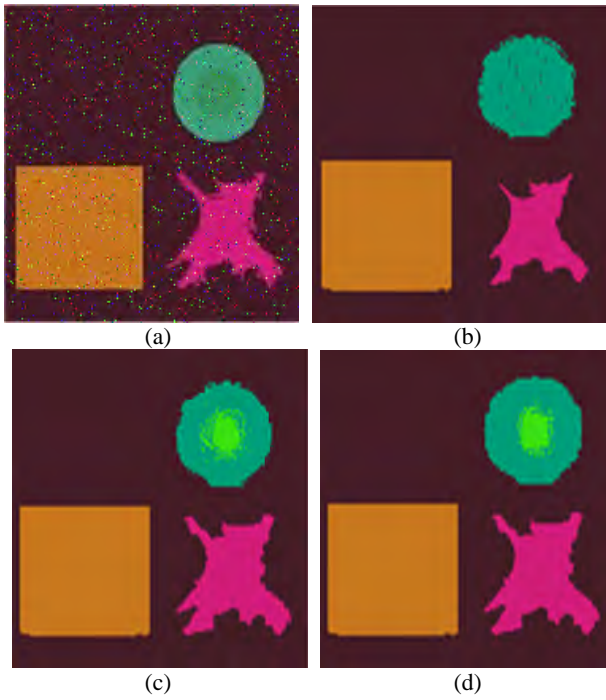


Figure 6. (a) Original image 256×256×3 color synthetic image disturbed with a "salt and pepper" noise and with gray level zero to 255 of each primitive colors and 6 classes, (b) Resulting image by HCM method, (c) Resulting image by FCM method, (d) Resulting image by the proposed method.

We notice that while using the proposed method, the regions are exactly and homogeneously segmented (see Figure 4(f)).

To evaluate the performance of the proposed segmentation algorithm, its accuracy was recorded.

Regarding the accuracy, Table 1 lists the segmentation sensitivity of the different methods for the data set used in the experiment.

The segmentation sensitivity [30] [31] (Sens %) is computed using

$$Sens = \frac{Np_{cc}}{N \times M} \times 100 \quad (21)$$

where: $Sens$, Np_{cc} and $N \times M$ correspond to the segmentation sensitivity (%), number of correctly classified pixels and the image sizes, respectively.

To provide insights into the proposed method, we have compared the performance of the proposed method with those of the corresponding Hard [29] and Fuzzy C-Means [5] algorithms. The method was also tested on synthetic images and compared with other existing methods.

The comparison of the proposed approach will be presented through the next experiment. Figure 6(b), (c) and (d) show the final segmentation results obtained from the HCM algorithm, the FCM and the TSMO-32 algorithms, respectively, when a "salt and pepper" noise of D density is added to the original image I , shown in Figure 6(a). This affects approximately ($D \times (N \times M)$) pixels. The value of D is 0.02.

Since HCM and FCM-based methods do not consider the spatial dependencies among the pixels and the compatibilities of the points belonging to the classes to compute the membership's degree, there are some isolated pixels that are remaining. As seen in the

experimental results, the proposed approach works better than the HCM and FCM-based methods.

Comparing Figures 6(b), 6(c), and 6(d), we observe that the five regions are correctly segmented in Figure 6(d), showing the good selection of the optimal thresholds by TSMO-32 method. It can be seen from Table 1 that 28.44% and 17.03% of the pixels were incorrectly segmented by HCM and FCM based methods, respectively, but only 04.68% are incorrectly segmented pixels by our proposed method. Comparing Figure 6(b) and 6(c) with 6(d), we can see that the image resulting from the proposed method is much clearer than the one resulting from the HCM and FCM based methods.

Also, for comparison, we test the performance of the proposed method in RGB and HSI color spaces by applying TSMO-32 to red, green, blue and hue color features, respectively (see Figure 3). The major problem that RGB color space suffers is a high correlation of the three color features. Also, the segmentation results obtained by using the color feature RGB is performed on three dimensions, therefore it needs a triple processing time. In fact, using HSI color space can solve this problem. The color feature Hue can represent subjective color suitable for human observes. The result does not have correlation and the segmentation is performed on only one dimension, using less processing time.

This could be observed in our experimental results. In Figures 3(b), 3(c) and 3(d), the different regions are misclassified during the segmentation using RGB space. Therefore, the same regions are segmented correctly using hue feature (see Figure 4(f)). This proves that hue is less influenced by shadow and highlight in an image.

Consequently, the Otsu method requires a lot of computation time, when extended to a multi-level threshold problem due to the fact that a large number of iterations are required for computing the cumulative probability and the mean of a class. In the present study, the TSMO-32 method [25] is used to greatly reduce the iterations required for computing the between-class variance in an image, and to determine the optimal thresholds. In fact, for a bi-level threshold ($M=2$) of an image, the iterations for the conventional Otsu's method [20], the TFCM method [27], and the TSMO-32 are 253, 167 and 36, respectively. Although, for ($M=5$) of an image, the iterations for the conventional Otsu's method [20], the TFCM method [27], and the TSMO-32 [25] are 250^4 , 225^4 and 26^4 , respectively. Hence, the iterations of the conventional Otsu's method [20] and the TFCM [27] are high compared to the iterations required by the TSMO-32 method.

5. Conclusion

In this paper, we have proposed a new method to color image segmentation based on multi-level thresholding technique. We extend the general idea of Otsu's method to multi-level thresholding in a medical color images. In the first phase of the segmentation, the histogram of an image is divided into groups which contain the same number of gray levels. In the second stage, the separate search range of Otsu's method should fall into the groups found in stage one. Hence, the optimal threshold for each



group can be found by the new fast algorithm called the TSMO method (Two-Stage Multi-level Otsu method). Experimental results are provided to demonstrate the effectiveness of the proposed method, and then a comparative study versus existing techniques is presented. The advantages of different color spaces, HSI and RGB, are also given. The color feature hue is proved to be more efficient than RGB color features by this research. The proposed method will find wide applications in image thresholding, segmentation, texture analysis, etc.

6. References

- [1] M. Wirth, D. Nikitenko and J. Lyon, "Segmentation of the Breast Region in Mammograms using a Rule-Based Fuzzy Reasoning Algorithm", *GVIP Journal Special Issue on Mammograms*, pp. 13-21, 2007.
- [2] N. R. Pal and S. Pal, "A review on image segmentation techniques" *pattern recognit.*, vol. 26, pp. 1277-1294, 1993.
- [3] P. K. Sahoo, S. Soltani, and A. K. C. Wong "SURVEY: A survey of thresholding techniques" *Comput. Vis. Graph. Image process.* vol 41, pp. 233-260, 1988.
- [4] R. E. Cummings, P. Pouliquen, and M. A. Lewis, "A Vision Chip for Colour Segmentation and Pattern Matching", *EURASIP Journal on Applied Signal Processing*, no. 7, pp. 703-712, 2003.
- [5] Y. Yang, C. Zhen and P. Lin, "Fuzzy C-means clustering algorithm with a novel penalty term for image segmentation", *Opto-Electronics Rev.*, vol. 13, no. 4, 2005.
- [6] M. B. Meevathi and K. Rajesh, "Volterra Filter for color image segmentation", *Inter. Journal of Computer Science and Engineering*, vol. 2, no. 1, 2008.
- [7] H. D. Cheng and Y. Sun, "A hierarchical approach to color image segmentation using homogeneity", *IEEE Transaction on Image Processing*, vol. 9, no. 12, pp. 2071-2082, 2000.
- [8] H. Seddik and E. Ben Braiek "Color Medical Images Watermarking, Based Neural Network Segmentation "GVIP Journal Special Special Issue on (Medical Image Processing), pp. 81-86, 2006.
- [9] H. D. Cheng, X. H. Jiang and J. Wang, "Color image segmentation based on homogram thresholding and region merging", *Pattern Recognition*, pp. 373-393, 2002.
- [10] S. Ben Chaabane, M. Sayadi, F. Fnaiech and E. Brassart, "Colour Image Segmentation using Homogeneity method and Data Fusion Techniques", *Eurasip journal on advances in signal processing*, 2009.
- [11] X. Gao, K. Hong, P. Passmore, L. Podladchikova, and D. Shaposhnikov, "Colour Vision Model-Based Method for Segmentation of Traffic Signs", *EURASIP Journal on Image and Video Processing*, October 2008.
- [12] E. Littmann and H. Ritter, "Adaptive colour segmentation – a comparison of neural and statistical methods", *IEEE Trans. Neural Network*, vol. 8, no. 1, pp. 175-185, 1997.
- [13] P. W. M. Tsang and W. H. Tsang, "Edge detection on object color", *IEEE Intern. Conf. On Image Processing-C*, pp. 1049-1052, 1996.
- [14] P. K. Saha, J. K. Udupa, "Optimum Image thresholding via class uncertainty and region homogeneity", *IEEE Trans. Pattern Anal. Mach. Intell.* 23 (7), pp. 689-706, 2001.
- [15] D. Aiteanu, D. Ristic, A. Graser, "Content based threshold adaptation for image processing in industrial application," In: *Internat. Conf. Control and Automation*, Budapest, Hungary, June, pp. 1022-1027, 2005.
- [16] H. F. Ng, "Automatic thresholding for defect detection", *Pattern Recognition Lett.* 27 (14), pp. 1644-1649, 2006.
- [17] F. Yan, H. Zhang, C. R. Kube, "A multistage adaptive thresholding method", *Pattern Recognition Lett.* 26 (8), pp. 1183-1191, 2005.
- [18] G. Jing, D. Rajan, C. E. Siong, "Motion detection with adaptive background and dynamic thresholds", *IEEE Inter. Conf. Information, Communications and Signal Processing*, Bangkok, Thailand, December, pp. 41-45, 2005.
- [19] C. Su, A. Amer, "A real-time adaptive thresholding for video change detection", *IEEE Internat. Conf. Image Processing*, Atlanta, Georgia, USA, October, pp. 157-160, 2006.
- [20] N. Otsu, "A threshold selection method from gray-level histograms", *IEEE Trans. Systems Man Cybern.* 9, pp. 62-66, 1979.
- [21] M. Portes de Albuquerque, I. A. Esquef, A. R. Gesualdi Mello, M. Portes de Albuquerque, "Image thresholding using Tsallis entropy", *Pattern Recognition Lett.* 25, pp. 1059-1065, 2004.
- [22] P. S. Liao, T. S. Chen, P. C. Chung, "A fast algorithm for multi-level thresholding", *J. Inf. Sci. Eng.* 17 (5), pp. 713-727, 2001.
- [23] S. K. S. Fan, Y. Lin, "A multi-level thresholding approach using a hybrid optimal estimation algorithm", *Pattern Recognition Lett.* 28 (5), pp. 662-669, 2007.
- [24] M. K. Quweider, J. D. Scargle, B. Jackson, "Grey level reduction for segmentation, thresholding and binarisation of images based on optimal partitioning on an interval", *IET Image Process.* 1 (2), pp. 103-111, 2007.
- [25] Deng-Yuan Huang and Chia-Hung Wang, "Optimal multi-level thresholding using a two-stage Otsu optimization approach", *Pattern Recognition Letters* 30, pp. 275-284, 2009.
- [26] W. Kim and R. Park, "Color image palette construction based on the HSI color system for minimizing the reconstruction error", in *IEEE Int. Conf. Image Processing*, pp. 1041-1044, 1996.
- [27] Y. W. Lim and S. H. Leung, "On the color image segmentation algorithm based on the thresholding and the fuzzy c-means techniques", *Pattern recognition*, pp. 935-952, 1990.
- [28] A. W. C. Liew and W. H. Lau, "fuzzy image clustering incorporating spatial continuity", *IEE Proceedings on Vision Image Signal Processings*, pp. 185-192, 2000.
- [29] R. Duda and P. Hart, "Pattern Classification and Scene Analysis", New York, Wiley, 1973.
- [30] V. Grau, A. U. J. Mewes, M. Alcaniz, R. Kikinis, and S. K. Warfield, "Improved watershed transform for medical image segmentation using prior information", *IEEE Transactions on Medical Imaging*, vol. 23, no. 4, pp. 447-458, 2004.
- [31] R. O. Duda, P. E. Hart, and D. G. Sork, "Pattern Classification", Wiley-Interscience, New York, NY, USA, 2000.



Biographies



Rafika Harrabi born in 1981 in Kairouan (Tunisia), she received the BSc degree in Electrical Engineering and the DEA degree in Automatic and Signal Processing from the High school of sciences and techniques of Tunis, respectively in 2007 and 2009. Currently, she is in the last

preparation year of its PhD. Its research interests are focused on signal Processing, image processing, classification, segmentation and data fusion.



Ezzedine Ben Braiek obtained his HDR on 2008 in Electrical engineering from ENSET Tunisia. He is, presently, professor in the department of electrical engineering at the technical university ESSTT and manager of the research group on vision and image processing at the CEREP. His

fields of interest include automatics, electronics, control, computer vision, image processing and its application in handwritten data recognition.





Multi-view Video Segmentation based on Bayesian Estimation and Graph Cut

Anh Tu Tran, Koichi Harada

Graduated School of Engineering, Hiroshima University, Higashihiroshima, Hiroshima city, 739-8521, Japan
[d103714, hrd]@hiroshima.u-ac.jp

Abstract

In this paper, we propose a method, which requires no interactive operation, to segment human object from multi-view video. Our method consist of two stages: for initial frame of the video sequence, we automatically extract object based on saliency model and iterated Graph cut. After having segmented object in first frame, we propose the algorithm combining Bayesian estimation and minimizing energy function using graph cut to segment object. In our energy function, the color, depth and spatial-temporal coherence are integrated in data term. Smooth term is encoded the penalty cost of the neighboring pixels with different labels. By combining Bayesian estimation, minimizing energy functions via graph cut is speeded-up. Experiment results on test sequences are encouraging.

Keywords: *Multi-views/Stereo Object segmentation, Automatic Object Segmentation, Object/Foreground extraction*

1. Introduction

Segmentation is one of the most common and active research topic in image processing image processing, computer vision for recently decades. Image/video segmentation can be classified into region-based segmentation and object-based segmentation. From the perspective of this study, video object segmentation can be classified into two main categories: segmentation in mono view and multi-view video.

Most of interest has been focused on the research of mono-view segmentation leading to many advanced algorithms, theories and technologies. Especially object-based segmentation has drawn great attention from the research and industrial community, resulting in many commercial products with image cutout tools or user interface, such as Video SnapCut [1], Lazy Snapping [2], Ratio cut [3] and GrabCut [4]. These interactive object segmentation tools requires user intervention to provide foreground/background hints by brush stokes or bounding box, which can obtain high accuracy object regions.

On the other hand, multi-view segmentation had not attracted much attention due to the limitation of capture technology. With the recent growing capability of

capturing devices, multi-view capture system with dense or sparse camera array can be built with ease, which motivates the development of multi-view techniques and its related applications. One of challenging task but playing very important roles in many in multi-view applications (image-based rendering, 3D reconstruction...) is object segmentation. Based on the different methodologies involved, the existing algorithms of object detection and extraction from multi-view images can be grouped into three categories: background modeling, visual hull-based and depth-based.

In this paper, we focus on depth-based multi-view object segmentation. Depth information recovered from multi-view image/video serves as an important cue for our segmentation algorithm. However, due to ill-posed nature of depth estimation, errors may occur in the depth map. To obtain more robust segmentation results for object-level manipulation, integration of depth, color, and other image cues should be considered. Depth reconstruction and multi-view segmentation is generally addressed in the sequential, joint or iterative fashion in a number of literatures. In [5], author described models and algorithms for bi-layer segmentation of stereoscopic frames. Stereo disparity is obtained by dynamic programming in Layered Dynamic Programming algorithm, and stereo match likelihood is then probabilistically fused with contrast-sensitive color model to segment each frame by ternary graph cut. In [6], multi-view scene reconstruction and segmentation are dealt with by joint graph-cut optimization for the challenging outdoor environments with moving cameras, for example, rugby and soccer scenes. Segmentation and depth labeling field are formulated into the unified energy function, which involves color and contrast term for segmentation, as well as the match and smoothness term for reconstruction. In [7], authors proposed a flexible and homogenous approach to simultaneous depth estimation and background subtraction in a multi-view setting, assisted by a static background image with known depth for each camera. The results of depth reconstruction and background separations algorithm is obtained as minimization of energy functional, to generation a dense depth map and foreground map. In [8], the estimated depth map and segmentation mask are



iteratively computed using an Expectation-Maximization (EM) algorithm.

In this paper, we approach in straightforward way. First, depth is estimated and then the depth information is incorporated into the segmentation framework. This approach is reasonable because depth maps are becoming a readily available commodity of the multi-view pipeline. The purpose of this research is to fuse color, contrast with depth to robust object segmentation. The contribution point of our research is automatically created trimap by Bayesian estimation. Created trimap is initial value to speed-up graph cut optimization algorithm. The remainder of the paper is organized as follows: Section (2) briefly introduces background of graph cut theory. Section (3) emphasizes on our proposal segmentation algorithm. Experiments and conclusion are presented in section (4) and (5) respectively.

2. Background of graph cut theory

Before going to further, we briefly explain the graph cut theory.

Graph cut based methods construct a graph topology to minimize the specified energy function activated by the max-flow/min-cut algorithm [8], so that the min-cut on the graph is of minimal energy among all the cuts separating the terminals.

Theoretically, $G = \{V, E\}$ is a directed graph with weighted edge capability. V is the vertices set including all the nodes as well as another two special terminals usually called sources and sink t . Generally, nodes represent pixels (or voxels), and source and sink correspond to the labels assigned to the nodes. E is a directed edge set, which connects all the vertices in the graph. It contains two types of edges in the graph named $t-link$ and $n-link$. $t-link$ connects pixels with terminals or labels. $n-link$ joins the pixels with their neighbors. All the edges in the graph are assigned weights or costs representing the edge capability. A $S-T$ cut C with two terminals is a partition of nodes in the graph into two disjoint sets S and T so that $s \in S$ and $t \in T$.

In combinatorial optimization, the cost of a cut is calculated as the sum of the edge weights passing from S to T . The min-cut problem is to find a cut with the minimal cost. In [10], they implemented a fast algorithm for graph-cut optimization with public source code, which drives the extensive application of graph cut technique in various optimization problems.

The general formulation of the energy function is defined as follows:

$$E(f) = \sum_{p \in P} E_p(f_p) + \lambda \sum_{(p,q) \in N} E_{p,q}(f_p, f_q) \quad (1)$$

where, f is the labeling field, P is the set of pixels, and N is the pixel's neighborhood system. $E_p(f_p)$ is called the data term and corresponds to the $t-link$ in the graph. It measures the likelihood of a certain pixel p assigned to the label f_p and how well the label f_p fits the pixel p given the observed data. The smoothness $E_{p,q}(f_p, f_q)$ is represented by the $n-link$ in the graph. It evaluates the

penalty of disagreement between p and q which are assigned with f_p and f_q respectively. λ is a parameter to weigh the importance of these two terms.

3. Proposal segmentation algorithm

3.1. Algorithm overview

Figure 1 shows the entire algorithm. Our algorithm takes the color image of key view, depth image which is estimated from the multi-views as the inputs, and the segmentation result of the foreground region as the outputs.

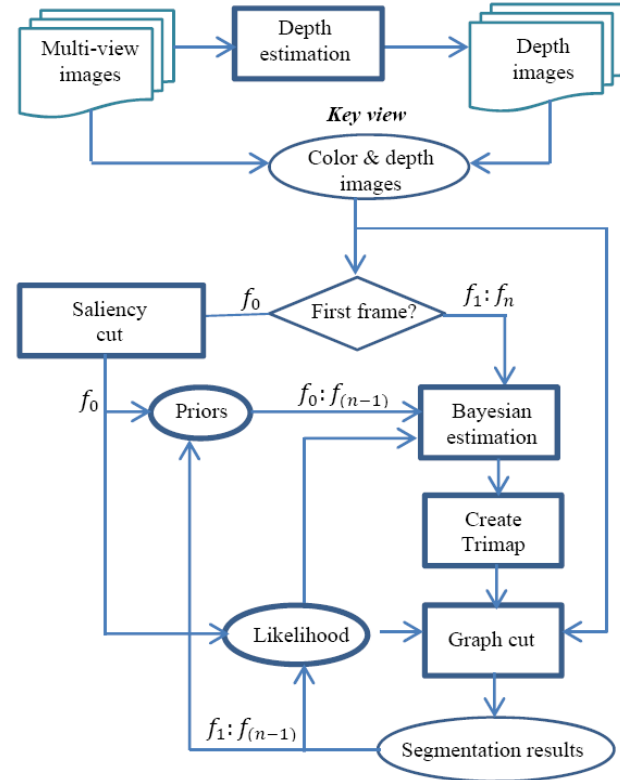


Figure 1. The illustration of proposal algorithm

In our framework, depth is estimated based on algorithms in [11], detail in section 3.2. But in many cases, depths are free given with multi-view colors images.

For the starting step, if the input is first frames in sequences, we apply the novel method to segment object based Saliency model [16] and GrabCut [4], given detail in section 3.3. From the second frame, we propose a probabilistic model combining Bayesian estimation and graph cut algorithm, discussion in section 3.4.

3.2. Depth estimation from multi-view images

Depth estimation aims at calculating the structure and depth of objects in a scene from a set of multiple views or images. This topic has been attracted extensive attentions in research communities. A comprehensive survey and evaluation of dense two-view stereo matching algorithms can be found in [12]. For multi-view scene reconstruction, one of the top advance algorithms is using graph cut and is described detail in series of papers [10],[11],[12].

In this work, depth is estimated based on algorithms proposal in [11]. In [11], the data term enforces photo-consistency. Let I be a set of pair of "nearby" 3D points. These points will come from different view, but they will



share the same depth (i.e., the points are of the form $\langle p, f_p \rangle$, $\langle q, f_q \rangle$ where $f_p = f_q$ and p, q are pixels from different views). Then the data term is:

$$\sum_{\{(p,l),(q,l) \in I\}} D(p,q)T[f_p = f_q = l] \quad (2)$$

The smooth term, on the other hand, involves a single camera at a time. It is defined to be:

$$\sum_{\{p,q\} \in N} D_{p,q}(f_p, f_q) \quad (3)$$

where, N is a neighborhood system on pixels in a single view.

The energy function in [11] is minimizing by expansion move algorithm [13], which is efficient approximation graph cut algorithms. Figure 2 shows an example of our depth estimation results.



Figure 2. Stereo depth image

Depth map provides vital information for scene interpretation; therefore, depth maps are becoming a readily available commodity of the multi-view pipeline. We can make use of this new free information to our algorithm.

3.3. Saliency cut in the first frame

Most of graph cut based object segmentation algorithm need user's intervention to specify the initial foreground and background regions as the hard constraints such as in [2], [4] and [15]. User's interaction is helpful to obtained good segmentation results, but the initialization itself may be annoying the user especially when much guidance is needed.

In our proposal algorithm, the object will be automatically segmented but requiring only the object mask in the first frame in video sequence. To obtain the first frame object mask, we have two ways: manually or automatically locating and extracting object. For manually extracting objects, user can use some tools such as Lazy Snapping [2], GrabCut [4], or using simple background subtraction method with having background of the first frame.

In this section, our purpose is automatically locating and extracting object in the first frame. Visual attention concept gives us with smart mechanism to perceptually attract human's attention toward the location of interesting objects in a complicated scene. Saliency model in [16] is one of the earliest works. Give a static image, this model employs color, intensity and orientation to compute Saliency Map (SM), which encodes the obviousness at each location in the visual input. Until now, there are many saliency models have proposed such as in [17], [18], [19], [20].

In this paper, we apply the ideal given in [17] and [18] to compute the saliency map. The process is demonstrated in Figure 3.

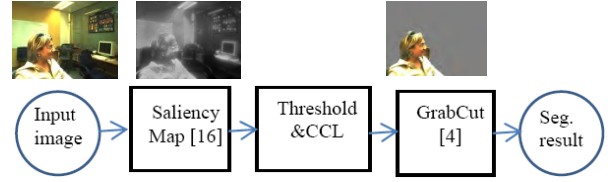


Figure 3. Object segment based on SM and GrabCut

After having saliency map, we consider the use of this map to assist in salient object segmentation. Saliency maps have been previously employed for unsupervised object segmentation such as in [20], [21], [22].

In our approach, we iteratively apply GrabCut [4] to refine the segmentation result initially obtained by threshold, dilation and erosion operating and connected component labeling (CCL) on the saliency map (see Figure 3). Instead of manually selecting a rectangular region to initialize the process, as in classical GrabCut, we automatically define the bounding region based on the result of threshold and CCL on SM.

Once initialized, we iteratively run GrabCut to improve the saliency cut result. We apply dilation and erosion operations on the current segmentation result to get a new "seed" for the next GrabCut iteration. Our experiments need 3 or 4 iterations to obtain good result.

Our approach can be the semi-automatic object segment by given some guidance to GrabCut as in classical one when saliency model have failed to locate the desire object.

3.4. Probabilistic model combining Bayesian estimation and graph cut

3.4.1. Color and depth data models

This session, we briefly explain the color and depth probability model using in our algorithm.

Color data model:

Gaussian Mixture Model (GMM) in RGB color space is used for color data model. We use two GMM models with K component, one for background and one for foreground. The likelihood of pixel p with color $C_p(r, g, b)$ (r, g, b : color values) belongs to the foreground or background can be written as:

$$P(C_p | f) = \sum_{k=1}^K w_{k,f} G(C_p; \mu_{f,k}, \Sigma_{f,k}) \quad (4)$$

where, $f \in \{F, B\}$ representing foreground and background; $w_{k,f}$ is a mixture weighting coefficient; and

$G(C_p; \mu_{f,k}, \Sigma_{f,k})$ is the k^{th} Gaussian component as:

$$G(C_p; \mu_{f,k}, \Sigma_{f,k}) = \frac{1}{(2\pi)^{3/2} |\Sigma_{f,k}|^{1/2}} \exp(-(C_p - \mu_{f,k})^T (\Sigma_{f,k})^{-1} (C_p - \mu_{f,k})) \quad (5)$$

Given the GMM model $\theta = \{w_{f,k}, \mu_{f,k}, \Sigma_{f,k}, f = (F, B), k = 1, \dots, K\}$ (i.e. the weight w , means μ , covariance Σ , and $2K$ Gaussian component for background and foreground), we can calculate the likelihood $P(C_p | f)$ with $f = \{F, B\}$ by using Equation (4).



Depth data model:

Depth image is an array of grey values. Here we use histogram of grey values for depth data model $h(p; f)$. The same as color model, we need two histograms, one for foreground and another for background. Histograms are normalized to sum to 1 over grey level range $\int_p h(p; f) = 1$ and we get likelihood of pixel p with depth value D_p as:

$$P(D_p | f) = h(p; f) \quad \text{with } f = \{F, B\} \quad (6)$$

3.4.2. Bayesian estimation and trimap creation

This section computes the probability of each pixel to be in foreground base on Bayesian estimation and the results are used to create the trimap, which is used for segmentation object via graph cut.

Lets C_p^t, D_p^t are color and depth value of pixel p on color and depth images at time t . The probability of pixel p belongs to foreground is calculated based on Bayes' formula as:

$$P^t(F | C_p, D_p) = \frac{P^t(C_p, D_p | F) P^t(F)}{P^t(C_p, D_p | F) P^t(F) + P^t(C_p, D_p | B) P^t(B)} \quad (7)$$

where, $P(x)$ is probability of x , F and B stand for Foreground and Background in which pixel p belongs to. We assume that color and depth are independent, so the likelihood:

$$\begin{aligned} P^t(C_p, D_p | F) &= P^t(C_p | F) P^t(D_p | F) \\ P^t(C_p, D_p | B) &= P^t(C_p | B) P^t(D_p | B) \end{aligned} \quad (8)$$

The likelihood $P^t(C_p | F)$ and $P^t(C_p | B)$ are calculated from foreground and background Gaussian Mixed Model (GMM) which are constructed from the previous segment result of color frame at $(t-1)$.

Similarly, the likelihood $P^t(D_p | F)$ and $P^t(D_p | B)$ are calculated from grey-level histogram which also constructed from previous segment result of depth image at $(t-1)$.

Because of successive frames in the temporal domain would have strong correlations, so the prior probability $P^t(F)$ and $P^t(B)$ of frame at t are calculated from the previous image frame at $(t-1)$. In order to get more accurate result, $P^t(F)$ and $P^t(B)$ can inferred from smooth map the 2D mask of segmentation results in previous frame at $(t-1)$ by performing Gaussian filter. Based on computed prior probability and likelihood probability, the posterior probability $P^t(F | C_p, D_p)$ and $P^t(B | C_p, D_p)$ are calculated by Equation (7).

Applying this process for whole image pixel p , we get the probability image $I_{prob}(p)$. Based on probability

value of pixel p , $I_{prob}(p)$ the trimap $T = \{F : \text{Foreground}; B : \text{Background}; U : \text{Uncertainly}\}$ can be created by:

$$\begin{cases} I_{prob}(p) \leq \text{threshold} & T(p) = B \\ I_{prob}(p) > 1 - \text{threshold} & T(p) = F \\ \text{otherwise} & T(p) = U \end{cases} \quad (9)$$

Here, threshold can be very small real value, in our experiment we choose threshold = 0.005. To remove the noise in trimap T , a filter is applied to foreground region and background regions.



Figure 4. Trimap (back color: background; white color: foreground; grey color: uncertainly)

3.4.3. Object segmentation by graph cut

In this part, we focus on setting the data term and smooth term for our algorithm.

Data term $E_p(f_p)$: measures the agreement between the segmentation labels and observer data. We define this term as the weighted sum in following equation:

$$E_p(f_p) = \alpha \sum_{p \in P} \{ -\log(P(C_p | f_p)) - \log(P(D_p | f_p)) \} + (1 - \alpha) \sum_{p \in P} -\log(I_{prob}(p)) \quad (10)$$

where, the likelihood $P(C_p | f_p)$ and $P(D_p | f_p)$ can be obtained from GMM for color cue and histogram for depth cues as describing in section 3.4.1. Different from others previous related works, we add new term $I_{prob}(p)$ as in the second row of Equation (10). This term encodes spatial temporal coherence to improve the segmentation result and can be reused the probability image in the creating trimap step.

Smooth term $E_{p,q}(f_p, f_q)$: measures the penalty of two neighboring pixel p and q with different labels and is defined as follow:

$$E_{p,q}(f_p, f_q) = \gamma \sum_{(p,q) \in N} T[f_p \neq f_q] \frac{1}{\|dist(p, q)\|} \exp(-\beta \|C_p - C_q\|) \quad (11)$$

where,

$$T[f_p \neq f_q] = \begin{cases} 1 & \text{if } f_p \neq f_q \\ 0 & \text{if } f_p = f_q \end{cases} \quad (12)$$

$\|disp(p, q)\|$ and $\|C_p - C_q\|$ are Euclidean distance of neighboring pixels in coordinate and color space respectively. N is the set of pairs of neighboring pixels. In practical experiment, a good results are obtained by defining pixels to be neighbors in 8-way connectivity (horizontal, vertical and diagonally). In [15], they had shown that it is more effective to set $\beta > 0$ since this



relaxes the tendency to smoothness in region of high contrast. We choose β the same in [15] as follow:

$$\beta = \left(2 \left\langle (C_p - C_q)^2 \right\rangle \right)^{-1} \quad (13)$$

This choice of β ensures that the exponential term in Equation (11) switches appropriately between high and low contrast. Note that if $\beta = 0$, the smooth term is well-known Ising model, which encourages smoothness everywhere [4].

In our algorithm, the automatically generated trimap $T = \{T_F, T_B, T_U\}$ (see section 3.4.2) is set as the initial values for graph cut. This is different point of our algorithms comparing to interactive segmentation method such as GrabCut [4], which requires considerable degree of user interaction for supplying trimap. With our trimap, Graph cut optimization only need perform in uncertainly region T_U . However, T_U is smaller region than whole image, so the time for graph cut optimization decreases considerably.

4. Experimental Results

To evaluate the performance of our method, we compare the segmentation results of our method with respect to the ground truth in the IU sequence, which can be freely downloaded from [24]. We define the Absolute Mean Error Rate (AMRE) of every fifth frame (in the left view) as the number of misclassified pixels over the total number of pixels in the image, which is the same measurement adopted in [5].

$$AMRE = \frac{1}{5} \sum_{i=0}^4 \frac{\text{misclassified pixel w.r.t ground trust}}{\text{total number of pixels in images}} \quad (14)$$

We make a comparison of two cases: 1) using only color/contrast cues, 2) fusing both color and depth cues. The comparison shows in Figure 5.

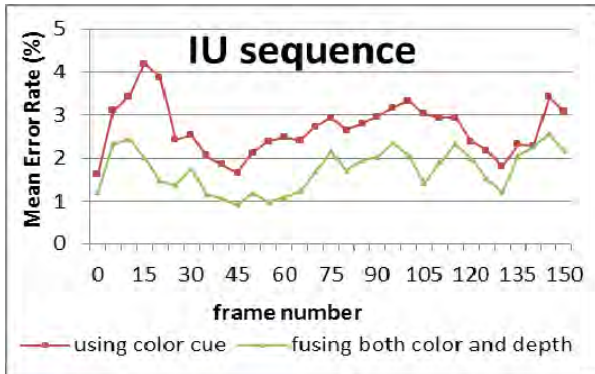


Figure 5. Quantitative comparison

Using only color/contrast is a simply our algorithm in which the depth is switched off. As in Figure 5, when depth and color are fused we can get the better segmentation results. The segmentation in the case using both depth and color is more stable than only color/contrast case because absolute mean error rate is not much varying.

Many frames in IU sequence, the background is non-stationary (there are other people moving in the background), but our algorithm is also working well on these frame, as demonstration in Figure 6.

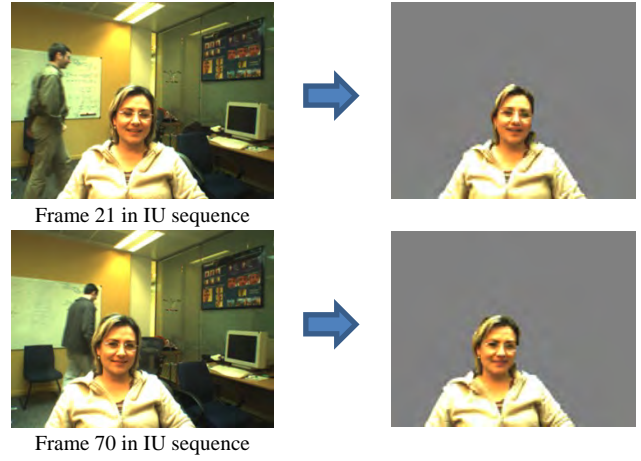


Figure 6. Segmentation works well with non-stationary background

Besides using the IU sequence given in [24], we also set up the system to capture stereo video to show result of our algorithm. A common Minoru 3D webcam [25] is used to capture stereo video. Bouguet's algorithm built in OpenCV [26] is used for camera calibration. Segmentation in our captured video also shows good results as in Figure 7.

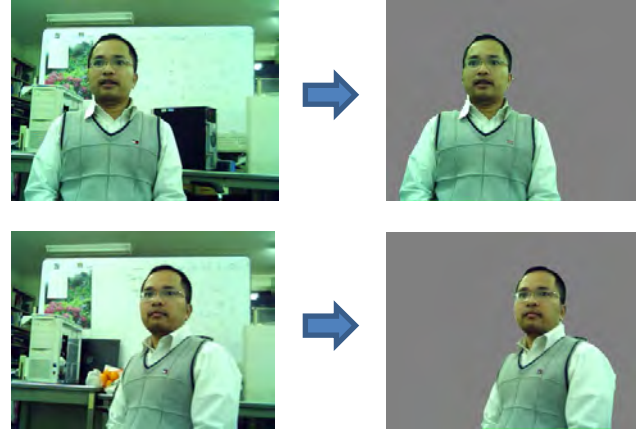


Figure 7. Segmentation on stereo video captured by Minoru 3D webcam

We are not focusing on optimizing the running time, but we measure the time which Graph cut process is initialized by our trimap comparing with the case the graph cut algorithm is applied over whole image. Our proposal is faster, about 100 milliseconds for each frame. All our implementations use C++ and OpenCV [26].

5. Conclusion

In this paper, we have introduced complete framework to automatically segment the human region in multi-view video. Depth is estimated from multi-view image by graph cut. Saliency model and iterated Graph cut are used to automatically locate and extract the interested object in the first frame to trigger the whole process. Depth is fused with color and spatial-temporal coherence in energy function. By combining Bayesian estimation, trimap is created automatically and used as initial value to speed-up minimizing energy functions via graph cut. Experiment results on various test sequences are encouraging. Our future work will be focused on improved this algorithm for multi-objects segmentation and efficient way to project the segmentation result in key-view to another views.



6. References

- [1] X. Bai, J. Wang, D. Simons, and G. Sapiro. Video snapcut: Robust video object cutout using localized classifiers. *ACM Transactions on Graphics*, vol.28(3), pages 19–31, 2009
- [2] Y. Li, J. Sun, C.K. Tang, and H.Y. Shum. Lazy snapping. *ACM Transactions on Graphics*, vol.23, pages 303–308, 2004.
- [3] S. Wang, J. M. Siskind. Image segmentation with ratio cut. *IEEE Transactions on PAMI*, vol.25(6), pages 675–690, 2003.
- [4] C. Rother, V. Kolmogorov, and A. Blake. GrabCut: Interactive foreground extraction using iterated graph cuts. *ACM Transactions on Graphics*, vol.23(3), pages 309–314, 2004.
- [5] V. Kolmogorov, A. Criminisi, A. Blake, G. Cross, and C. Rother. Probabilistic fusion of stereo with color and contrast for bilayer segmentation. *IEEE Transactions on PAMI*, vol. 28(9), pages 1480–1492, 2006.
- [6] J.Y. Guillemot, J. Kilner, and A. Hilton. Robust graph-cut scene segmentation and reconstruction for free-viewpoint video of complex dynamic scenes. In *Proc. ICCV*, pages 809–816, Japan, 2009
- [7] B. Goldlücke, M.A. Magnor. Joint 3d-reconstruction and background removal separation in multiple views using graph cuts. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, page 683–688, 2003
- [8] N. Grammalidis, L. Bleris, and M.G. Strintzis. Using the expectation-maximization algorithm for depth estimation and segmentation of multi-view images. In *Inter. Sym. on 3D Data Processing Visualization and Transmission*, pages 686–689, 2002.
- [9] L. Ford, D. Fulkerson. *Flows in network*. Princeton University Press, 1962.
- [10] Yuri Boykov, Vladimir Kolmogorov. An Experimental Comparison of Min-Cut/Max-Flow Algorithms for Energy Minimization in Vision. In *IEEE Transactions on PAMI*, vol. 26(9), pages 1124–1137, 2004.
- [11] V. Kolmogorov, R. Zabih. Multi-camera scene reconstruction via graph cuts. In *European Conference on Computer Vision*, vol.3, pages 82–96, 2002.
- [12] D. Scharstein, R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision*, vol. 47(1-3), pages 7–42, 2002.
- [13] Boykov, Veksler, Zabih. Fast Approximate Energy Minimization via Graph Cuts. *IEEE Transactions on PAMI*, vol. 23(11), pp. 1222–1239, 2001
- [14] Kolmogorov, Zabih. What Energy Functions can be minimized via Graph Cuts? *IEEE Transactions on PAMI*, vol. 26(2), pages 65–81, 2004.
- [15] Y. Boykov, M.P. Jolly. Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images. In *Proc. Int. Conf. on Computer Vision*, vol. 1, pages 105–112, 2001.
- [16] L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on PAMI*, vol.20(11), pages 1254–1259, 1998.
- [17] Lingyun Zhang, Matthew H. Tong, Timothy K. Marks, Honghao Shan, and Garrison W. Cottrell. SUN: A Bayesian Framework for Saliency Using Natural Statistics. *Journal of Vision*, vol.8(7):32, pages 1–20, 2008.
- [18] Lingyun Zhang, Matthew H. Tong, Garrison W. Cottrell. SUNDAY: Saliency Using Natural Statistics for Dynamic Analysis of Scenes. In *Proceedings of the 31st Annual Cognitive Science Conference*, Amsterdam, Netherlands, 2009.
- [19] R. Achanta, S. Hemami, F. Estrada, and S. Sussstrunk. Frequency-tuned salient region detection. In *CVPR*, pages 1597–1604, 2009.
- [20] Y.-F. Ma, H.-J. Zhang. Contrast-based image attention analysis by using fuzzy growing. In *ACM Multimedia*, pages 374–381, 2003.
- [21] B. Ko, J. Nam. Object-of-interest image segmentation based on human attention and semantic region clustering. *Journal of Optical Society of America A*, vol. 23(10), pages 2462–2470, 2006.
- [22] J. Han, K. Ngan, M. Li, and H. Zhang. Unsupervised extraction of visual attention objects in color images. *IEEE TCSV*, vol.16(1), pages 141–145, 2006.
- [23] M.A.A. Medeiros, L.A.D.S. Cruz. Iterative disparity estimation and image segmentation. *Proceedings of the 8th conference on Signal processing, computational geometry and artificial vision*, pages 67–72, Rhodes, Greece, 2008
- [24] <http://research.microsoft.com/en-us/projects/i2i/data.aspx>.
- [25] Minoru 3D webcam, <http://www.minoru3d.com/>
- [26] Opencv, <http://opencv.willowgarage.com/wiki/>

Biographies



Anh Tu Tran is a PhD candidate of the Graduate School of Engineering at Hiroshima University. He received the B.E in 2001 from Hanoi University of Technology, Vietnam, and received the ME in 2007 from RWTH Aachen University, Germany. His current research is mainly in area of computer vision and image processing.



Koichi Harada is a professor of the Graduate School of Engineering at Hiroshima University. He received the BE in 1973 from Hiroshima University, and MS and PhD in 1975 and 1978, respectively, from Tokyo Institute of Technology. His current research is mainly in the area of computer graphics. Special interests include man-machine interface through graphics; 3D data input techniques, data conversion between 2D and 3D geometry, effective interactive usage of curved surfaces. He is a member of ACM, IPS of Japan, and IEICE of Japan.





Accurate Detection of Microaneurisms and Hemorrhage for the Planning of Laser Treatment

Nahed Solouma, Ph.D. and Amro El-Dib, MD.

National Institute of Laser Enhanced Sciences (NILES), Cairo University, Giza – Egypt

nsolouma@niles.edu.eg, dramreldib@yahoo.com

www.niles.cu.edu.eg

Abstract

Retinal disorders are among the most dangerous diseases because many of them may lead to blindness if not early diagnosed and managed. Medical screening is very important as it allows the early diagnosis of the disease. Only ophthalmologists who have enough experience can differentiate between normal and abnormal retinas at the early stages of the disease. So, screening routines of the retina are very expensive and rarely to be done. In this work we provide a computer-aided screening system for the retinal disorders. The system could be used easily by the young physicians as it can automatically detect the early symptoms of abnormalities such as microaneurysms, hemorrhage and exudates. Using newly proposed image processing algorithms, we implemented a screening method based on the experience of the ophthalmology experts.¹

Keywords: *Computer-aided diagnosis, hemorrhage, microaneurysm detection.*

1. Introduction

Blindness is an intimidating problem. Many life difficulties and risks could be caused by blindness. According to the World Health Organization (WHO), blindness is defined as a corrected visual acuity below 3/60 on the Snellen scale for the best eye, or a central visual field diameter of less than 10°. It is estimated that about 180 million people worldwide have a visual impairment. Of these, 45 million persons are blind and 135 million are partially sighted. These figures are expected to double over the next 25 years due to combination of an increasing population and aging worldwide. The most common causes of blindness are cataract, glaucoma, and retinal disorders such as age-related macular degeneration (AMD) and diabetic retinopathy (DR) [1]. While cataract and glaucoma can be treated surgically, there is no curative treatment for the

macular degeneration and the retinal vascular diseases till now. So, AMD and retinal vascular diseases, specially diabetic retinopathy, are the major causes of untreatable blindness worldwide. To prevent blindness caused by retinal disorders, they should be early diagnosed and then the spread of the disease can be stopped. So, screening of the retina is of most importance in this aspect. Unfortunately, screening of the retina is costly and should be achieved by highly qualified ophthalmologists. This motivated us to design and implement a computer-aided screening system that can automatically detect the abnormal symptoms on the retina in their early stages. This system could be used by young physician because it depends only on capturing digital images of the retina. These images can be processed by the system to decide whether the patient has the markers of retinal disorders, that might lead to blindness if left untreated, or not. The system is assumed to be able to determine the severity of the disease and the urgency for management.

Since AMD and (DR) are the major causes of blindness, we provide robust methods for the early detection of such diseases. In their early stages, these diseases are characterized by the presence of microaneurysms, hemorrhage and exudates in the retina. While exudates appear as yellow irregular regions in the fundus photograph, microaneurysms and hemorrhage appear as red spots with irregular boundaries for the later as shown in figure 1. According to how much these symptoms appear on the retina, the severity of the disease could be estimated. In a previous work, we published a new method for the detection of blood vessels and exudates in retinal images using morphological image processing techniques [2].

In this work we propose some methods for the detection of hemorrhage and microaneurysms together with the presentation of the screening system. The aim was originally to build a screening system that can be used easily for quick assessment of retinal diseases. Screening systems are very useful tools in medicine as they save time and efforts of medical experts. Young doctors always use screening systems for initial diagnosis and

¹ This study has been implemented on Matlab platform at Biomedical Engineering lab. NILES – Cairo university.



deciding whether the investigated patient needs more diagnosis and treatment or not.

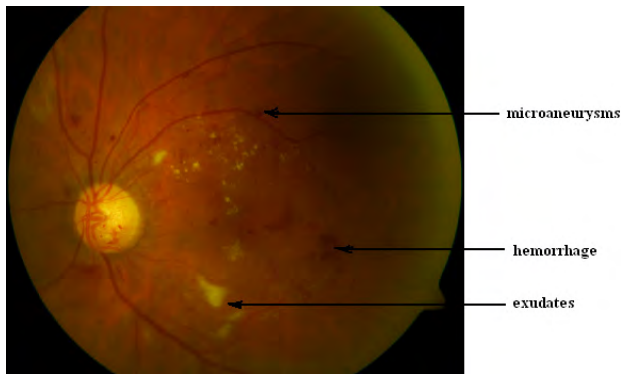


Figure 1. image of a retina with microaneurysms, hemorrhage and exudates pointed to.

Several techniques have been developed for the detection of hemorrhages and microaneurysms in fundus images. Joshi et al. in [3] used image subtraction to extract the blood vessels and hemorrhages. They also used mathematical morphology to suppress blood vessels and to highlight only hemorrhages. The authors in [4] used feature extraction based on pixel classification to detect red candidates and distinguish between microaneurysms and hemorrhages using k-nearest neighbor classifier. Region growing segmentation technique, intensity threshold and edge enhancement are used to extract red lesions are presented in [5]. Then they differentiate between hemorrhage and microaneurysms using neural network classification. Shade correction and image normalization of the green channel then diameter closing is used to detect candidates which are classified into real microaneurysms and other objects in [6]. Meindert et al. [7] compared the results of five different methods, produced by five different teams of researchers on the same set of data for microaneurysms detection. In one method, normalization of the green component of the retinal images using median filtering. Then, vessels are detected and removed using a top-hat transform and morphological reconstruction. The candidates are obtained and each candidate is segmented using region growing technique. A new approach based on multi-scale correlation filtering (MSCF) and dynamic threshold was developed in [8] to detect microaneurysms. Automated microaneurysm detection method based on double-ring Filter was achieved by the authors in [9]. In [10], a method to detect the hemorrhages using hue saturation value (HSV) space was proposed. Automated fundus photograph analysis algorithms for the detection of primary lesions and a computer-assisted diagnostic system for grading diabetic retinopathy (DR) and the risk of macular edema (ME) are introduced [11].

In this work, we propose a fast and an accurate method for the detection of microaneurysms and hemorrhages in fundus photographs. The detection of such disorders could be greatly improved by the new pre-processing algorithm which we propose in this paper. The power of the proposed pre-processing technique comes from the fact that it selectively enhances the contrast of microaneurysms and hemorrhage. Because the presence of microaneurysms and hemorrhage is the early signs of

diabetic retinopathy, their detection would be an important issue for screening purposes.

The of the paper is organized as follows: The materials and methods used in this work, are explained in details in the following section. The results are presented and discussed in the last section. Finally, we conclude our work and point to our future work.

2. Materials and Methods

A dataset of about one hundred retinal images were collected for this work. They include most of the retinal disorders as well as some images of normal retinas. Some of the image in this dataset were collected using a VISUCAM Carl Zeiss fundus camera located at the biophotonics lab – NILES – Cairo University. The others were collected from the free access database STARE.

Fundus images usually taken as colored or graylevel images. Often, The green component of the colored image has the highest contrast and more information because it's like red-free illumination images. Normally, the blood stream in the graylevel images appear as dark objects in brighter background. While in the dye-injected patients, they appear as bright objects in darker background which results in higher contrast because of the fluorescence effect. Trials are always done to minimize the use of the dye because of cost and patient comfort requirements. So, researchers try to increase the contrast of images taken without the injection of a dye. Also, it is difficult to the physician to capture a high quality fundus images because of many reasons such as non uniform illumination, the limited control of the patient movement and the focusing required from the patient and for the system itself. So, image pre-processing is always required to enhance the images.

Proposed method: In an attempt to overcome the above-mentioned drawback related to enhancement of retinal images, we tried many of the pre-processing techniques. As we are interested in detecting dark objects such as microaneurysms and hemorrhage, we proposed a new pre-processing technique that can selectively increase the contrast of such objects. Segmentation of the anatomical structures and lesions can then be achieved using suitable algorithms. The steps of the proposed method are described below.

Pre-processing: The images collected for this work were investigated to select the images with hemorrhage, microaneurysms and exudates. The images were then pre-processed to selectively increase the contrast of microaneurysms and hemorrhage because of the interest in detection of these lesions. We proposed a new preprocessing technique based on frequency domain filtering. As it is well known, if a high pass filter is applied to an image, it will sharpen this image. That is to say that the contrast of objects in relation to their backgrounds will be increased. So, high-pass filtering of an image increases the intensities of the objects while decreasing the intensities of their backgrounds as shown in figure 2 c. So, subtracting a high pass-filtered image from the original image would smooth the bright objects in relation to their backgrounds while decreasing the intensities of dark objects in relation to their backgrounds. Since we are interested in detecting microaneurysms and hemorrhage which appear as dark



objects in the graylevel images, we can use this filtering and subtraction technique to selectively increase the contrast of dark objects such as microaneurysms and hemorrhage while minimizing the contrast of bright objects such as exudates and optic disk. The pre-processing method can be described using the following equation:

$$g(x, y) = f(x, y) - af_{hpf}(x, y) \quad (1)$$

Where: $f(x, y)$ is the original graylevel image, $f_{hpf}(x, y)$ is the high pass-filtered image multiplied by a factor a and $g(x, y)$ is the resultant image. Figure 2 shows the original colored image and its green component in the upper row and the result of high pass filtering and the final result in the lower row. From this figure, we can observe that the contrast of the bright objects (exudates) is increased. This filtered image is then multiplied by 2 and subtracted from

the original graylevel image. This leads to decreasing the contrast of exudates and enhancing the appearance of microaneurysms and hemorrhage in the final image.

Segmentation of retinal images: A good screening system should be able to detect any type of abnormality associated with diabetic retinopathy with high sensitivity. The most important signs are microaneurysms, hemorrhages (red lesions) and exudates (bright lesions). Also it is important to extract the normal structures such as the blood vessels tree, the optic disc and the macula. In a previous work, we provided robust algorithms to detect the optic disc, the blood vessels, the macula and the exudates [2]. To get a good screening system, we provide some methods to detect the other two abnormalities: microaneurysms and hemorrhages.

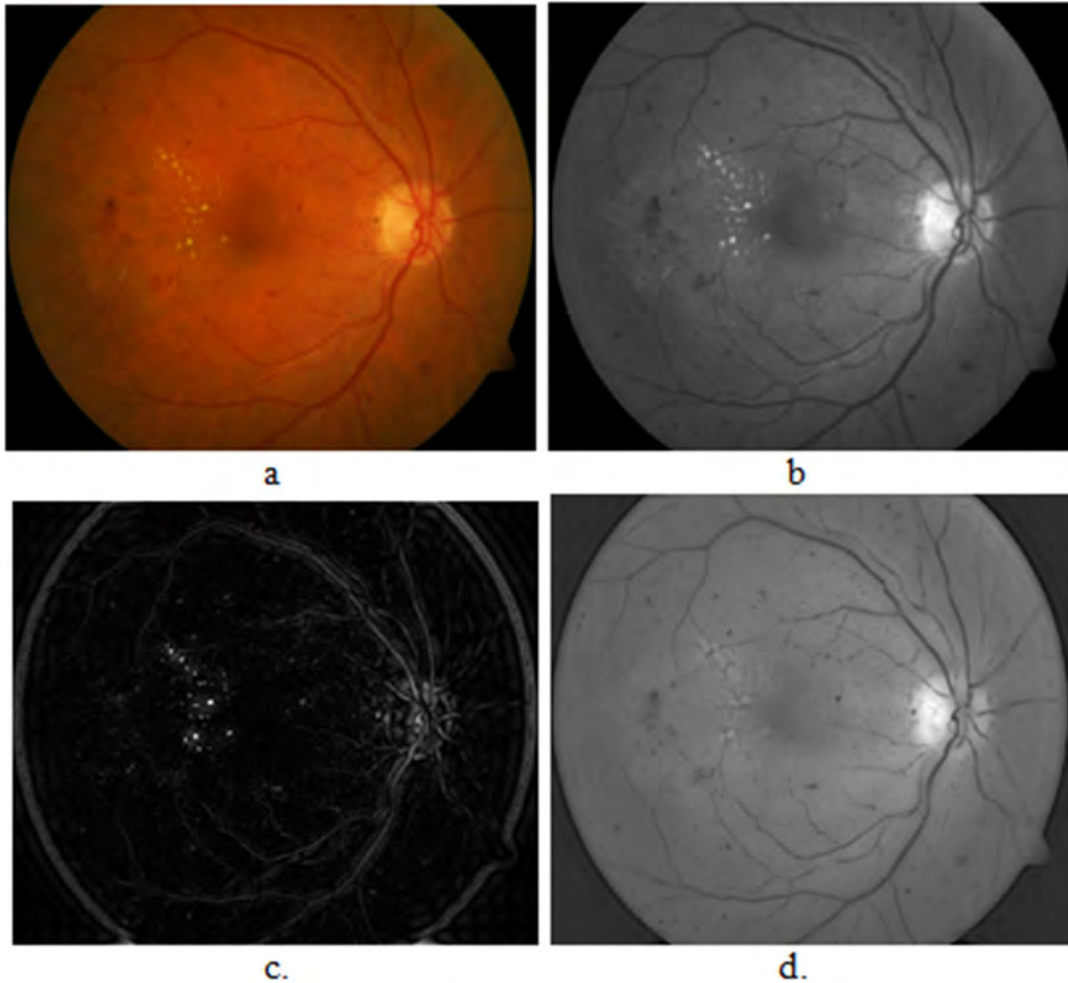


Figure 2. a. Original colored image, b. green component, c. high pass-filtered image and d. pre-processed image

Microaneurysms appear as red spots of circular shape with similar color as blood vessels in color fundus images. Their diameter is smaller than $125 \mu\text{m}$. They may appear as disconnected from the blood vessels tree because they are situated on capillaries, and capillaries are not visible in color fundus images [6],[12]. Also hemorrhages have the same color as the blood vessels but with irregular shapes. The difference in shape and size between microaneurysms and hemorrhages help us to distinguish between them.

After applying the pre-processing step described above, microaneurysms and hemorrhages will be of higher contrast. As microaneurysms and hemorrhages are dark regions in a bright background, we can detect them by using morphological filling operation on graylevel images. If the original is subtracted from the filled one, then we have an image that contain all the dark regions (blood vessels tree, macula region, microaneurysms and hemorrhages). We can obtain a binary image by simple thresholding as demonstrated in figure 3.



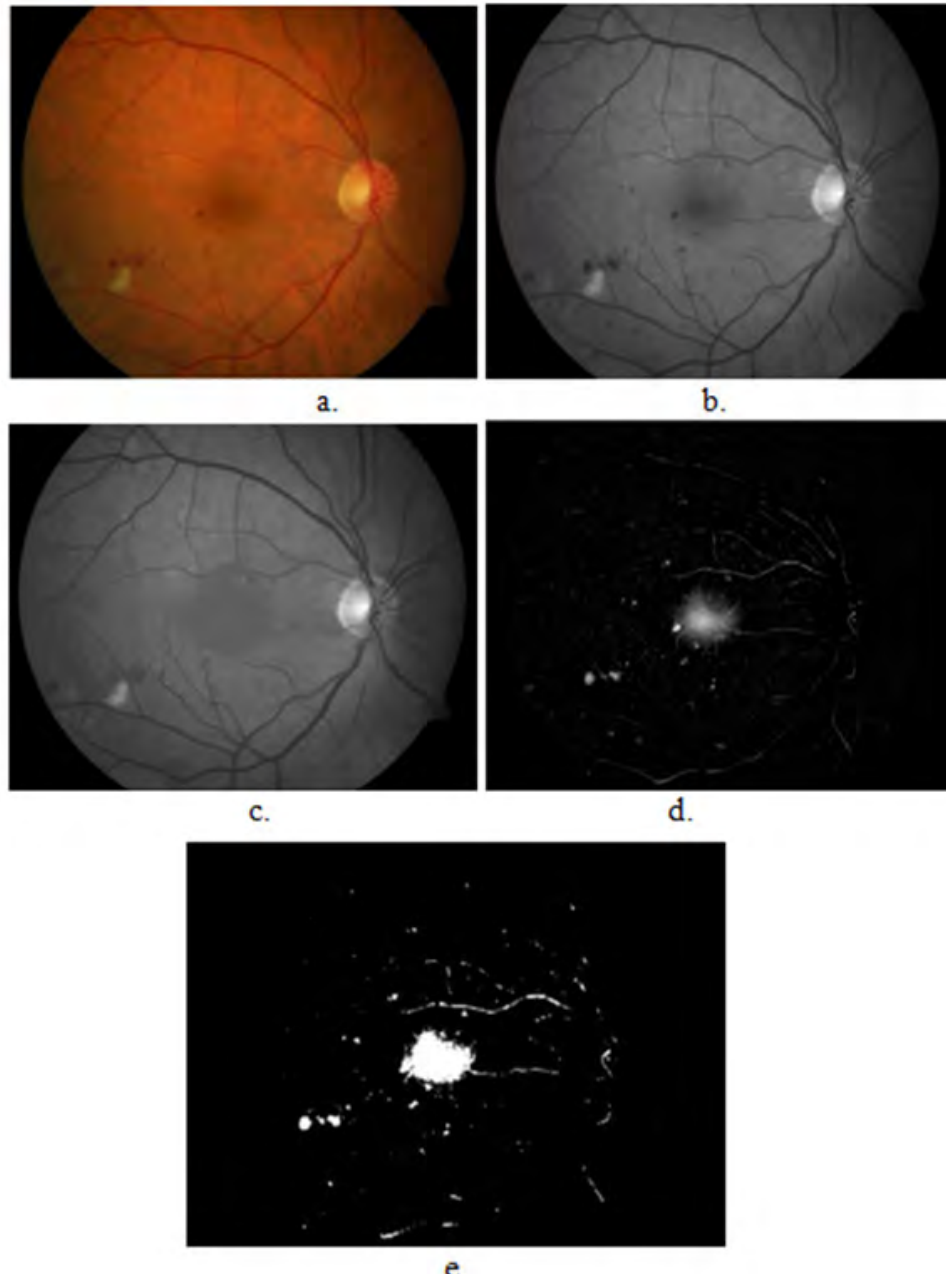


Figure 3. Dark regions detection: a. original image, b. green component, c. green component after filling operation, d. subtraction of c and b and e. dark regions after threshold.

Detection of blood vessel tree: Now, we need to extract the blood vessels from the result image to obtain an initial estimation of candidate regions (microaneurysms and haemorrhages). This could be done by closing the binary image using two line structuring elements of different sizes rotated by different angles as in described in [2]. The detected blood vessels is shown in figure 4 a.

Macula detection: The macula is the area of acute vision within the retina. It appears as dark and homogeneous area near the optic disc, approximately 2.5 times the diameter of the optic disc from the centre of optic disc [13]. It can be extracted by region growing algorithm. To segment the macula, we manually select a point in the fovea (the center of the macula), then by a region-growing algorithm, the macula was extracted as the connected region around the fovea as described in [14]. Figure 4 b shows the result of macula detection.



Figure 4. Initial estimation of microaneurysms and haemorrhages: a. blood vessels detection, b. macula detection and c. initial estimates of lesions.

Detection of microaneurysms and haemorrhages: After the removal of blood vessels and macula from the image obtained by morphological filling and segmentation, the result is an initial estimation of microaneurysms and haemorrhages as shown in figure 4 c. The final detection of these lesions can be obtained by applying the



morphological reconstruction algorithm described in [15, 16] and given by:

$$h_{k+1} = (h_k \oplus b) \cap I_{in} \quad (2)$$

Where, h_k is the marker image at the k th iteration (h_1 is the image contains the inverted initial estimate of microaneurysms and hemorrhages superimposed on the inverted original image), b is the structuring element and I_{in} is the input image (the inverted green component). This is an iterative process which must be repeated until no changes occur in h . The final iteration result is then subtracted from the inverted green component and thresholded to get the final estimate of microaneurysms and hemorrhages. Figure 5 shows the different steps of this method.

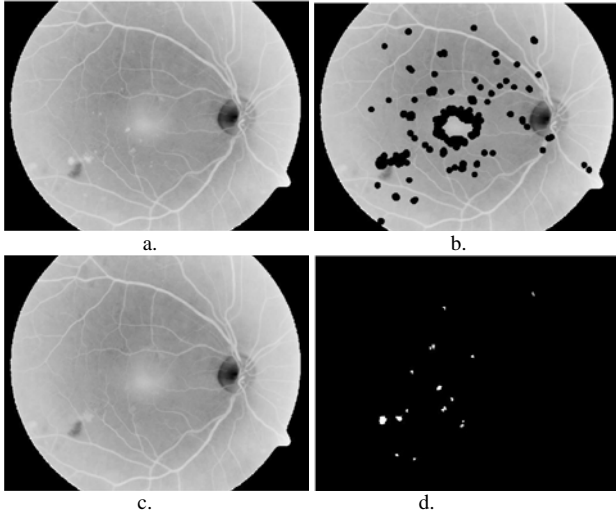


Figure 5. Image reconstruction: a. mask image I_{in} , b. marker image, c. the result of reconstruction process and d. final estimation of microaneurysms and hemorrhages after thresholding.

Differentiation between microaneurysms and hemorrhages: The obtained binary image contains a final estimation of microaneurysms and hemorrhages together. Now we need to separate the microaneurysms from hemorrhages and remove any false features come from noise. Microaneurysms are close to circular structures with small diameters but hemorrhages are irregular shapes. To separate them, all the connected components in the binary image are detected. For each connected components, the area, the minor axis length and the major axis length are to be determined to check the circularity of each connected component. A component is said to be circular if its c value given by equation 3 tends to 1.

$$c = \text{abs}\left(1 - \frac{\text{MajorAxisLength} - \text{MinorAxisLength}}{\max(\text{MajorAxisLength}, \text{MinorAxisLength})}\right) \quad (3)$$

The result of differentiating microaneurysms from hemorrhage is demonstrated in figure 6.

Results: As mentioned above, the aim of this work is to develop a retinal screening system that can help early diagnose of retinal disorders to prevent blindness. Major steps in such system is to provide some means by which the abnormalities can be detected. Since we are interested in detecting the retinal abnormalities that can lead to blindness, we implemented some algorithms that can be used to accurately detect the exudates, microaneurysms and hemorrhage as early signs of retinal disorders.

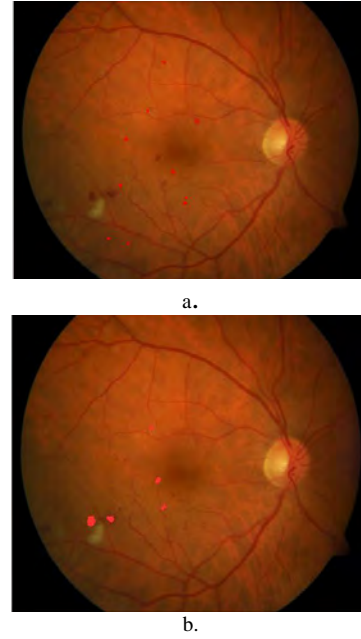


Figure 6. a: microaneurysms and b: hemorrhages.

A new pre-processing algorithm based on subtracting high pass-filtered of the green component from the original one is proposed. The application of this algorithm to retinal images, enhances the appearance of the dark spots which represent the hemorrhage and microaneurysms in fundus photographs. As demonstrated in figure 2, the result of this pre-processing algorithm increases the accuracy of red lesion detection. This algorithm has been tried on our dataset and the best results of red lesion enhancement were obtained using values of 1 and 2 for the factor a which is the multiplication of the high-pass filtered image.

A combination of the morphological algorithms used for the detection of hemorrhage and microaneurysms is used in this work to obtain more accurate results than that obtained previously. This resulted in increasing the sensitivity of red lesion detection in the fundus photographs. Together with our previously published exudates detection algorithms, these algorithms are used in our proposed screening system. The system provides a computer program with a user-friendly interface that help physician in diagnosis and treatment planning. The system can capture fundus images from cameras, process them and save them for further analysis and study. It also provides general image processing technique as enhancement and montage synthesis. A snapshot of the screening system is shown in figure 7. In this figure, the captured image can appear in the input frame. Many images from a video sequence can be taken and put into a montage to get the full extent of the human retina for better diagnosis and treatment planning. At the lower left of the figure, the processed retinal image are displayed. The first demonstrates the blood vessel tree as this is of great importance for the ophthalmologists. In the second image, the blood vessel tree, the optic disc and the macula are segmented for further investigation that can be done by the physicians. Lesions are detected and presented in different colors in the third image. Exudates appear as yellow structures where hemorrhage and microaneurysms are presented at different degrees of red color.



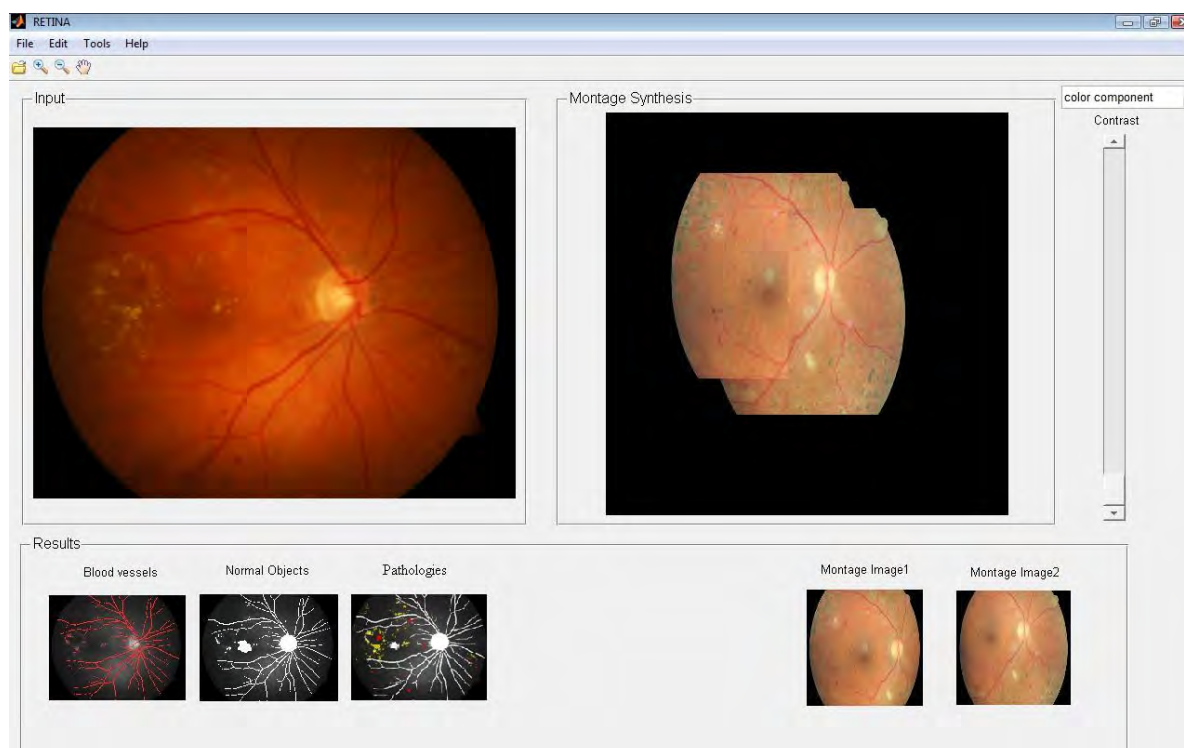


Figure 7. a snap shot of the screening system.

3. Conclusion

In conclusion, we provided a new image pre-processing techniques to selectively improve the detection of microaneurisms and hemorrhage as early signs of diabetic retinopathy. The technique is based on sharpening the image using high-pass filtering. Then, selective smoothing of the bright objects in relation to their background while increasing the contrast of the dark objects. This could be efficiently used to improve the detection of microaneurisms and hemorrhage in retinal images using common segmentation algorithms. The proposed method together with our previous work used to detect exudates in retinal images can be used to implement a screening system for the retinal diseases.

This is the conclusion. The major contributions of this article arise from the formulation of a new approach, to the modelling and identification of XYZ and ABC features that provide improved computational efficiency in the positioning techniques. By manipulating the manner in which feature information of both A and BVN signatures is incorporated into the model, it can be shown that significant improvements in the performance of the algorithm can be realised. Moreover, the simplicity and the efficiency of dynamic pose tracking techniques succeeded to improve the robot pose estimation process.

4. Acknowledgements

This work is funded by the Egypt EU Innovation Fund (EEIF) and the RDI program of the ministry of higher education and scientific research of Egypt. We should also thank Engineer Doaa Youssef for her great assistance.

5. References

- [1] Members of monitoring committee (2006) State of world's sight vision 2020: the right to sight, Blindness, World Health Organization WHO. Pp 18.
- [2] Doaa Youssef, Nahed Solouma, Amr El-dib, Mai Mabrouk, and Abo-Bakr Youssef. (2010) New Feature-Based Detection of Blood Vessels and Exudates in Color Fundus Images. Image Processing Theory, Tools and Applications 978-1-4244-7249-9/10/\$26.00 ©2010 IEEE explore.
- [3] Joshi Manisha Shivaram, Dr.Rekha Patil, Dr. Aravind H. (2009) Automated Detection and Quantification of Haemorrhages in Diabetic Retinopathy Images Using Image Arithmetic and Mathematical Morphology Methods. International Journal of Recent Trends in Engineering, Vol 2, No. 6.
- [4] Niemeijer M., Van Ginneken B, Staal J, Suttorp-Schulten MS, Abramoff MD. (2005) Automatic detection of red lesions in digital color fundus photographs. IEEE Trans Medical Imaging Vol.24, No 5, pp.584-592.
- [5] Usher D, Dumskyi M, Williamson TH, Nussey S. (2004) Automated detection of diabetic retinopathy in digital retinal images: a tool for diabetic retinopathy screening. Diabetic Med, Vol.21, pp.84-89.
- [6] Thomas Walter, Pascale Massin, Ali Erginay, Richard Ordonez, Clotilde Jeulin, Jean-Claude Klein. (2007) Automatic detection of microaneurysms in color fundus images. Medical Image Analysis, Vol. 11. 555–566.
- [7] Meindert Niemeijer, Bram van Ginneken, Michael J. Cree, Atsushi Mizutani, Gwénolé Quéllec, Clara I. Sánchez, Bob Zhang, Roberto Hornero,: Automatic Detection of Microaneurysms in Digital Color



Fundus Photographs. IEEE TRANSACTIONS ON MEDICAL IMAGING, VOL. 29, NO. 1.

- [8] Bob Zhang, Xiangqian Wu, Jane You, Qin Li, Fakhri Karray. (2010) Detection of microaneurysms using multi-scale correlation coefficients. Journal of Pattern Recognition vol. 43. 2237–2248.
- [9] Atsushi Mizutani, Chisako Muramatsu, Yuji Hatanaka, Shinsuke Suemori, Takeshi Hara, Hiroshi Fujita. (2009) Automated microaneurysm detection method based on double-ring filter in retinal fundus images. Medical Imaging Proc. of SPIE Vol. 7260 72601N-1.
- [10] Yuji Hatanaka, Toshiaki Nakagawa, Yoshinori Hayashi, Masakatsu Kakogawa, Akira Sawada, Kazuhide Kawase, Takeshi Hara, Hiroshi Fujita. (2009) Improvement of Automatic Hemorrhages Detection Methods using Brightness Correction on Fundus Images. Medical Imaging Proc. of SPIE Vol. 6915 69153E-1.
- [11] Dupas B, et al. (2010) Evaluation of automated fundus photograph analysis algorithms for detecting microaneurysms, haemorrhages and exudates, and of a computer-assisted diagnostic system for grading diabetic retinopathy. Diabetes Metab. doi:10.1016/j.diabet.2010.01.002.
- [12] Chandran Abraham, Annie Mathai. (2009) Diabetic retinopathy for the clinician. Jaypee brothers medical publishers (p) ltd, 1st ed.
- [13] Akara Sopharak, Bunyarit Uyyanonvara, Sarah Barman, Thomas H. Williamson. (2008) Automatic detection of diabetic retinopathy exudates from non-dilated retinal images using mathematical morphology methods. Journal of Computerized Medical Imaging and Graphics. Vol. 32. 720–727.
- [14] Mai S. Mabrouk, Nahed H. Solouma, Yasser M. Kadah. (2006) Survey of Retinal Image Segmentation and Registration. GVIP Journal, Vol. 6, Issue 2, pp 1:8.
- [15] Oille P. (2003) Morphological image analysis: principles and applications. 2nd ed., Heidelberg: Springer.
- [16] Rafael C. Gonzales, Richard E. Woods. (2002) Digital image processing. Second Edition by Prentice Hall.

Biographies



Author One received her B.Sc., M.Sc. and PhD from the department of biomedical engineering, faculty of engineering, Cairo university. She is currently an associate professor of biomedical engineering at the National Institute of Laser Enhanced Sciences (NILES), Cairo university. She is leading a group working in the area of ophthalmic research at the biophotonics labs at NILES.



Author Two received his B.Sc., M.Sc. and MD from the department of Ophthalmology, faculty of Medicine, Cairo university. He is currently an Assistant Professor at NILES Cairo University, in the biophotonics Laboratory. He is a Member of the ophthalmology clinic at NILES.







Optimal Watermarking Scheme in Time Frequency Domain for most Robustness and better Imperceptibility

Hassène Seddik and Ezzeddine Ben Braiek

ESSTT, 5 Av. Taha Hussein, 1008, Tunis, Tunisie

E-mails, hassene.seddik@esstt.rnu.tn and Ezzeddine.BenBraiek@esstt.rnu.tn

Abstract

The widespread of new image technologies and data exchanges has created the need for new techniques able to insure copyright protection and data owner identification. Considering the deficiencies of steganographie and cryptography, watermarking is nowadays a new technique that provides an efficient mean for resolving these problems. Watermark embedding techniques depend on the representation domain of the image (spatial, frequency, and multi-resolution). Every domain has its specific advantages and limitations. Moreover, each technique in a chosen domain is found to be robust to specific sets of attack types. The ultimate goal of each technique is to allow the embedded watermark to be invisible and robust against a wide range of attacks. The most damaging unintentional attacks are compression techniques. Many of these last techniques are based on the new coding procedures using Multi-resolution domain. In order to develop a watermarking technique robust against large sets of attacks, we must exploit this domain to overcome these attacks constraints. In this paper, a study is developed considering all the parameters contributing in the watermarking scheme. An optimal point allowing robustness and imperceptibility of the embedded watermark is established. Using these optimized parameters any technique based on the multi-resolution domain can reach the maximum of imperceptibility and robustness. These results can immediately be injected in previous algorithms to improve their results or to build new techniques outputting directly better results.

Keywords: *robust Image watermarking, wavelet transform, Stirmark attacks, parameters optimization.*

1. Introduction

The increased commercial activity on Internet and media industries, demand protection of media such as images, video and audio against illicit processing and use. Digital representation of copy-righted materials offers various advantages; however, the fact that an unlimited number of perfect copies can be illicitly produced is a serious threat to the right of content owners. [5, 14, 24, 16 and 10]. Watermarking is used for owner identification and authentication by verifying whether the data has been

used, changed or altered [15 and 17]. A research field whose results could undoubtedly help in the design of a watermarking algorithm is image compression. The image compression is based on removing from images all the data which are perceptually irrelevant with respect to the human system visual [1]. On the other hand, embedding an imperceptible watermark is made in such a way that it remains visually invisible and does not damage the image contents. As a matter of fact, the new generation of image coding such as standard JPEG2000, MPEG-4 [4 and 7], rely strongly on DWT to obtain good quality images at low coding rates. In order to avoid that the embedded watermark disappears when applying a compression to the image, the use of the same domain (DWT) for watermark embedding where the most used compression algorithms are based is strongly recommended [6]. This procedure allows following the steps of image coding and avoids inserting a watermark in the DWT coefficients that are removed after a compression is applied [2]. However, a major problem must be resolved. How can we choose the different parameters taking part in the watermarking procedure? In fact if one of these parameters is poorly selected it is difficult to reach the uppermost of imperceptibility and robustness of the embedded watermark [3 and 18]. In this paper, a study trying to optimize all the parameters intervening in the watermarking process is conducted. An equilibrium point generating the optimal parameters for the embedding procedure is established. These parameters are injected in the watermarking algorithm in order to reach the best equilibrium between the imperceptibility and robustness criteria. By the use of these parameters the algorithm is found to be robust against synchronous and asynchronous transforms. The attacks are generated by the STIRMARK tool [13, 25]. This paper is organized as the following: In section II an overview of watermarking techniques in the DWT domain is presented showing the state of art and the different parameters related to the watermarking process. The experimental configuration and the simulation results aiming to optimize all these parameters are presented in the third section. The privileges that this optimization offer and its efficiency when injected in a watermarking algorithm to increase robustness and maintain the imperceptibility under the visual threshold is summarized in the conclusion.



2. Watermarking in the DWT domain

1. Overview

DWT has gained interest among watermarking researchers because of its ability to model the HVS behavior. As witness to this domain importance, countless papers proposing different algorithms have been proposed over the last few years. Many of them take inspiration from the most used wavelet based compression algorithms. [9], [20] and [22]. Other algorithms hide into images binary logos which are also hierarchically decomposed [12]. A binary logo is decomposed by DWT, and added several times into the sub-bands of the DWT decomposition of the host image. In [8], the binary logo and the image are hierarchically decomposed through DWT; each detail sub-band of the logo is then embedded into the corresponding one of the image. This is a non blind technique. In [11], a binary watermark is coded in selected coefficients of the detail bands. In the extraction process the watermark is recovered by analyzing coefficients quantization, a threshold is fixed to decide about correlation value between the original watermark and the extracted one. An excellent tutorial on the possible approaches to exploit HVS characteristics for watermarking images and videos is presented in [21], a variable strength of the coded watermark on each DWT band according to HSV is developed. In [13] the authors selected the low pass band to insert the watermark. A difference is imposed among the mean values of two equally sized, randomly selected, subsets of the low pass image; a blind technique is presented.

In [19], the watermark is embedded using the LSB technique of DWT sub-band coefficients, the selected modified bits are chosen with respect to the HSV. As a substitutive technique watermark extraction is blind. In all these proposed techniques and in other ones the magnitude of the different parameters and coefficients used depends only on the embedding procedure and the image characteristics. The increase of each parameter is thresholded by the appearing distortions caused by the inserted watermark. But when several parameters intervene in the algorithm they may interfere. For this reason we must not chose for every parameter alone but the optimal values that lead to an equilibrium between them which allow us the maximum of robustness against attacks and the minimum of perceptibility that must remain under the psycho-visual invisibility threshold.

Three main and essential parameters are used in most DWT watermarking algorithm: the embedding strength factor, the decomposition level and the sub-band used for insertion. Many techniques try to fix one or two of these parameters and vary the last in order to maximise the robustness of the algorithm without damaging the watermarked image. This procedure will decrease the reliability of the proposed algorithm even if they are satisfying but less that we can get by a parameters optimisation [23]. In this paper all these variable parameters are tested and optimised through experimental results. The results evaluation is based on three essential criterions: the psycho-visual sensitivity, the PSNR factor and the resistance against synchronous and asynchronous attacks. Taking into the consideration the wavelet kinds, equilibrium between the cited criteria can be obtained

basing on this research in order to generate optimal watermarking shames.

2. Watermarks database

A watermarks database with 50 different watermarks is built up. These used watermarks are binary images with size equal to $P \times P$, with $P=50$ as shown by the following equation and figures:

$$W_L = \{W_L(i, j), 0 \leq i, j \leq P\}, \quad W \in \{0, 1\} \quad (1)$$

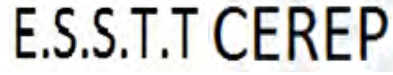


Figure 1: used watermarks

These watermarks includes personal and general information concerning the research unit name..etc.

3. Experimental configuration

Many experiments and tests are developed. The watermarking procedure is as follows: Once the original image is transformed in the time-frequency domain after applying lossless wavelet decomposition, the watermark is coded. After recovering the spatial representation of the watermarked image many STIRMARK attacks are applied in order to test the robustness of the embedding algorithm. These tests in addition to other distortions measurements allow as in each test checking and finding the equilibrium point between the involved parameters. The following equation presents the watermark inserting procedure:

$$f_w(i, j) = f_{mean} + (f(i, j) - f_{mean}) \cdot (1 + \alpha \cdot w_k) \quad (2)$$

$f_w(i, j)$ represents the watermarked image, f_{mean} is the mean of the used sub-band coefficients, α is the embedding strength and w_k is the inserted watermark.

This described procedure is repeated in each test. In every set of tests one parameter is varied, the other ones are fixed and the optimum is checked.

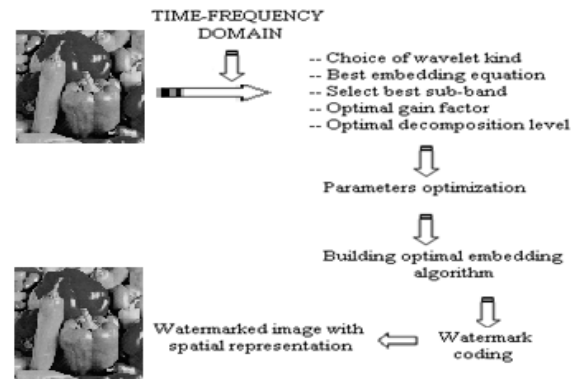


Figure 2: Steps of watermarking scheme

3. Experimental implementation and tests

In these experiments five steps are planned. In the first we chose the optimal wavelet that generates no loss in the decomposed signal and less distortion to the watermarked image. In the second step we determine in which sub-band of details the watermark is coded with respect to a better compatibility with the human system visual. Embedding in the sub-band that generates less distortions allow us to maximize the amount of coded data and increase of the embedding potency. The better embedding equation is fixed in the third step. The equation that engenders fewer



distortions with the same embedding force is chosen. This embedding force called gain factor are optimized in the forth step. In the last step, the wavelet decomposition level is optimized. Figure 2 describes theses steps.

1. Optimal wavelet choice

In time frequency domain, wavelet can generate loss or lossless decomposition. Furthermore, adding noise to the frequency sub-bands of a decomposed image provoke distortions in its spatial representation. But it is important to know that the potency of these distortions differs with the wavelet kind. In this section, we check witch wavelet type can affect the less the spatial image intensities when the watermark is added. The diagonal sub-band is used for watermark insertion and the gain factor coefficient is fixed as 0.8. The table below sum up the comparison study. After inserting the watermark, the PSNR between the original image and the watermarked one is computed. The wavelet decomposition that engenders the less distortion to the spatial representation of the watermarked image is considered.

	Wavelet Kind	PSNR (dB)
1	Haar Haar	37.3
2	Daubechie2 db 2	38.65
3	Daubechie4 db 4	38.85
4	Biorthogonale3.7 Bior 3.7	41.83
5	Symle4 Sym 4	41.5
6	Symlet8 Sym 8	41.89
7	Coiflet2 Coif 2	41.5

Table 1: Different wavelets used and the corresponding generated distortions.

Using the sixth wavelet, the watermarked image presents lower amount of introduced distortions.

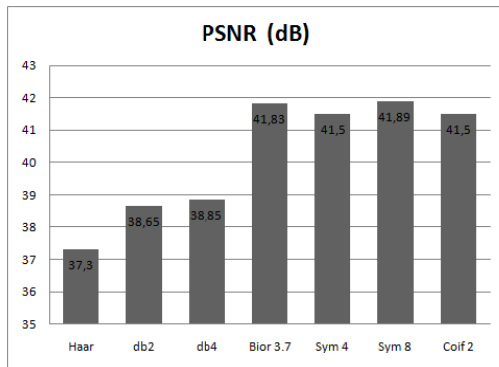


Figure.3: optimisation of the wavelet kinds

2. Optimal frequency sub-band choice

An image can be transformed by performing a DWT in both vertical and horizontal directions, resulting in one low frequency sub-band (LL) and three high frequency sub-bands (LH, HL, and HH).

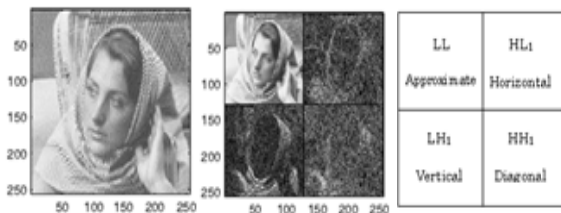


Figure 4: First level wavelet decomposition

In fact this decomposition causes some problem with respect to the selection of the sub-band in witch is better

coding the watermark. We don't know which one from the LH, HL or HH sub-bands agree the main with the HSV and cause fewer distortions to the original image if their coefficients are changed by the watermark coding. To answer this inquiry, a study is carried out. It consists in fixing the parameters that have not been optimized and insert the same watermark with the same embedding gain factor in the different frequency sub-band. The insertion that engenders lower distortion to the original image is considered. Consequently we can increase the gain factor and maximize the robustness of the algorithm. The watermark insertion in the different sub-bands is described by theses equations:

$$Y_{0,i,j}^{LH} = X_{0,i,j}^{LH} + (X_{0,i,j}^{LH} - X_{0,i,j}^{LH_{mean}})(1 + \alpha \cdot W_{iN+j}) \quad (3)$$

$$Y_{0,i,j}^{HL} = X_{0,i,j}^{HL} + (X_{0,i,j}^{HL} - X_{0,i,j}^{HL_{mean}})(1 + \alpha \cdot W_{MN+iN+j}) \quad (4)$$

$$Y_{0,i,j}^{HH} = X_{0,i,j}^{HH} + (X_{0,i,j}^{HH} - X_{0,i,j}^{HH_{mean}})(1 + \alpha \cdot W_{2MN+iN+j}) \quad (5)$$

As $Y_{0,i,j}$ is the watermarked sub-band, $X_{0,i,j}$ is the original sub-band mean, W is the watermark and α is the embedding strength. The result is evaluated by comparing the PSNR of the watermarked image in each sub-band. The table below presents the gathered results.

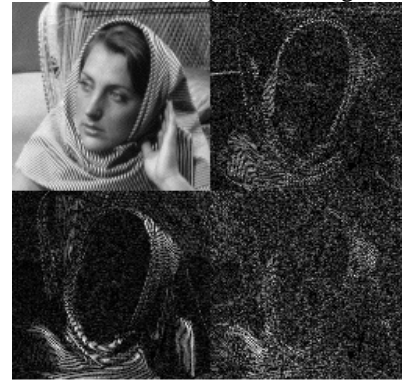


Figure.5: Watermarked cameraman sub-bands

Gain factor coefficient α	PSNR (HL1 sub-band insertion)	PSNR (LH1 sub-band insertion)	PSNR (HH1 sub-band insertion)
0.1	50.49	48.11	55.18
0.2	44.47	42.09	49.27
0.3	40.95	38.57	45.81
0.4	38.45	36.07	43.29
0.5	36.51	34.13	41.36
0.6	34.95	32.55	39.82
0.7	33.39	31.21	38.50
0.8	32.43	30.05	37.41
0.9	31.41	29.03	36.43

Table 2: The PSNR variation the respect to the used sub-band in the watermarking process

By comparing the results, we find that the HH sub-band presents more reliability with respect to the HSV and causes less distortion to the processed image if it is used to code the watermark in. This sub-band is fixed to be used in our watermarking process.

3. Selection of the optimal embedding force " α "

The embedding force or strength α called gain factor is the first parameter in charge of the robustness of any watermarking algorithm. The increase of this factor causes an increase of the watermark resistance to different synchronous and asynchronous attacks. The



limit of this factor increase is thresholded by the invisibility limit of the coded watermark. If this threshold is exceeded, distortions may appear in the watermarked image causing visual degradation, revealing the presence of an external sequence. The selection of this parameter must be well done in order to maximize the robustness of the watermark without exceeding the imperceptibility threshold. The criterion used to establish the limit of the gain factor increase and distortion appear is the PSNR. A limit of 37dB is fixed to decide about the presence of visible distortions in the watermarked image. The following figures illustrate the distortion appearing with the gain factor variation. More α increases more the watermarked image is distorted.

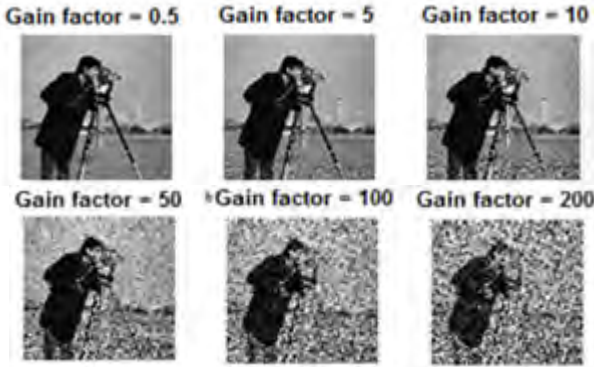


Figure 6: Impact of the gain factor variation

Gain factor α	PSNR (dB)
0.1	55.18
0.2	49.27
0.3	45.81
0.4	43.29
0.5	41.36
0.6	39.82
0.7	38.50
0.8	37.41
0.9	36.43

Table 3: PSNR progress with the gain factor variation

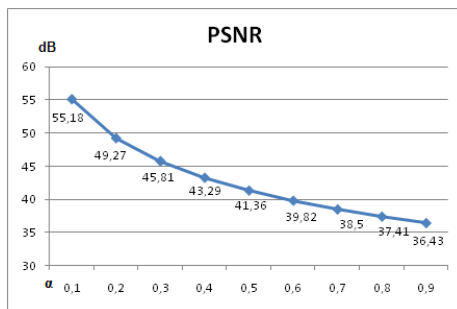


Figure 7: Distortion variation with the gain factor increase

As shown by table 3, the gain factor corresponding to the fixed PSNR threshold is $\alpha = 0.8$. This is the optimal embedding strength that guarantees the maximum robustness and minimum imperceptible distortions.

4. choice of the most useful embedding equation

Different equations are proposed in the literatures. The most used are the following:

$$f_w(i, j) = f_{mean} + (f(i, j) - f_{mean}) \cdot (1 + \alpha \cdot w_k) \quad (6)$$

$$f_w(i, j) = f(i, j) \cdot (1 + \alpha \cdot w_k) \quad (7)$$

To improve the watermarking algorithm in to the limits of robustness, we have to decide which equation leads to a more robust watermark coding. Both equations spread the watermark on the image border. The first adjust the intensity of α with the mean of the image variation. The second embed the watermark linearly in the sub-bands coefficients of the image.

Gain factor α	PSNR (dB)	Gain factor α	PSNR (dB)
0.1	55.50	0.6	41.20
0.2	49.7	0.7	39.80
0.3	46.00	0.8	38.70
0.4	43.7	0.9	37.4
0.5	41.80		

Table 4: PSNR progress with the gain factor variation using the second equation

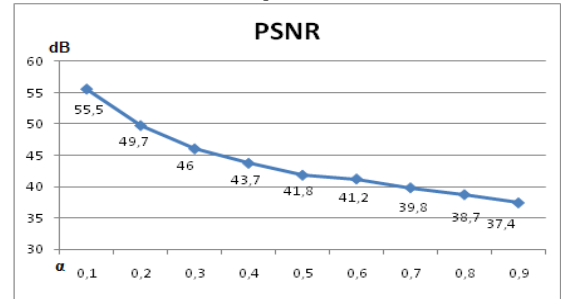


Figure 8: Distortion variation with the gain factor increase using the second embedding equation.

The second equation is more reliable and improves the robustness of the algorithm compared with the first one.

5. Selection of the optimal decomposition level

As described in the last section, the first decomposition generates three high frequency sub-bands (LH, HL, and HH). The same process is repeated on the LL sub-band to generate the next level of decomposition. For an n-level decomposition for an $M \times N$ image, the size of the area in which watermarks are to be embedded is $\frac{MN}{2^{2n}}$, as n is

the level number.

Decomposition level	PSNR (dB)
Level 1	39.83
Level 2	37.31
Level 3	35.88
Level 4	34.24
Level 5	34.38
Level 6	30.5

Table5: Distortion variation against level decomposition

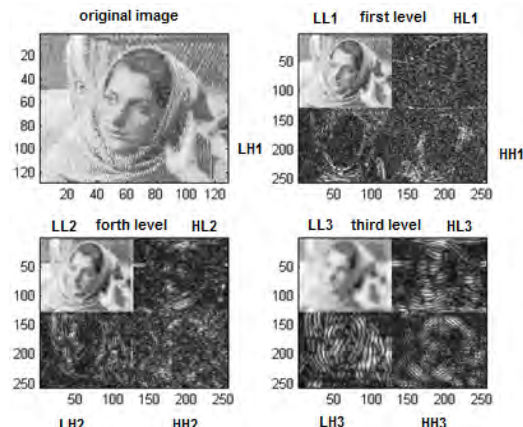


Figure 10: Third level decomposition of LENA image.



The figure above, show a LENA original image after its decomposition until the third level. The following figures present the original image and watermark tracked by the watermarked images in different decomposition levels and respectively the images illustrating the differences between the watermarked image and the original one. It is clear that the watermark is fixed on the borders of the image. From these figures representing the zones of the watermark coding we deduce that more the decomposition level is high, more the watermark is spread and distributed over and near the borders. This distribution is controlled in the frequency domain depending on the non-randomly selected sub-bands coefficients.

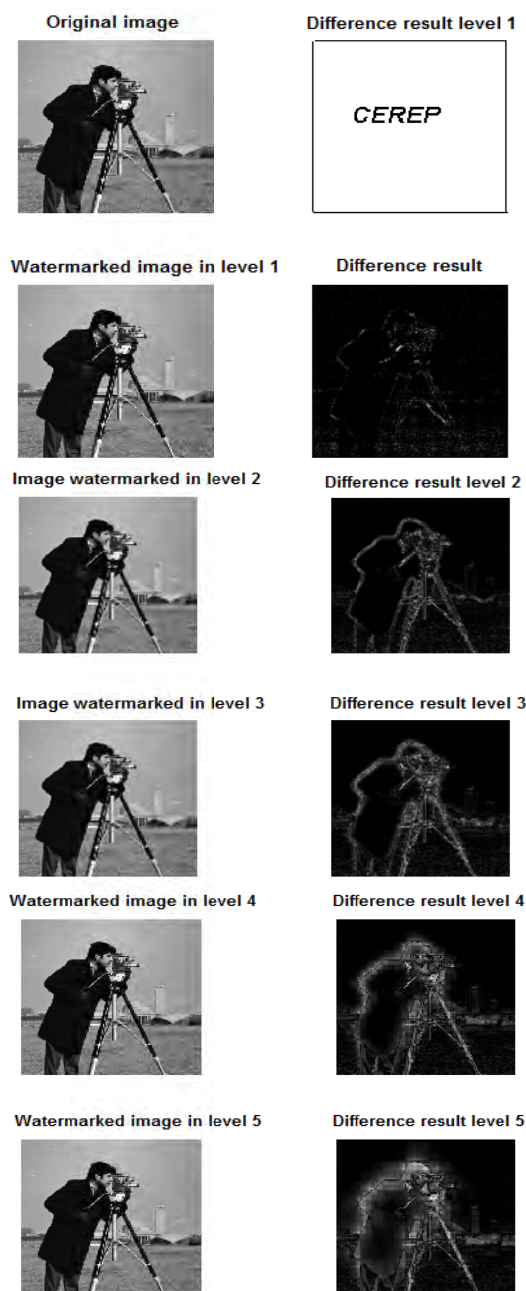


Figure 11: Different level watermarked sub-bands and the difference between the watermarked and original image

The following figure represents the PSNR variation with respect to the level decomposition. The Distortions introduced to the watermarked image are higher when the level decomposition of the image is superior

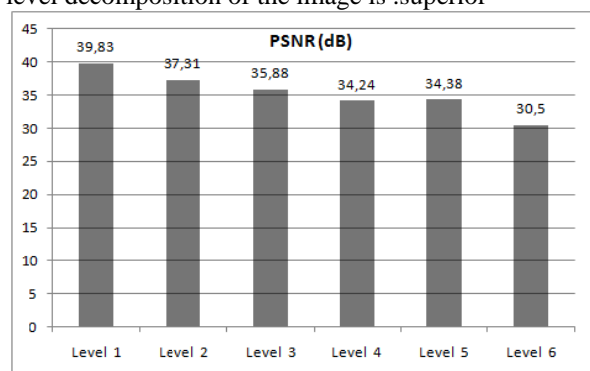


Figure 12: Distortion variation against level decomposition

The first level is found more advantageous for watermark embedding procedures. It generates fewer distortions when compared with levels of higher order.

4. CONCLUSION

In this paper, an overview over different watermarking techniques in the DWT domain is explored. Processing the different parameters that intervene in the watermarking process a strategy to optimize these parameters is built up. Every parameter is processed distinctly to the others to find its optimal value that leads to a robust and imperceptible watermarking algorithm. Experiments and tests are conducted to locate these values. Injecting these optimal parameters in the embedding equation of the watermarking process, guarantee better robustness of the watermarked image against different attacks and decreases the distortions to maintain them under the perceptibility threshold. This optimization wasn't done in the literature and can be exploited easily for DWT watermarking schemes.

5. REFERENCES

- [1] M. Barni, F. Bartolini and A. Piva, 'Improved wavelet based watermarking through pixel wise masking', IEEE Transaction on Image Processing, Vol. 10, no. 5, May 2001.
- [2] M. Barni, F. Bartolini, A. De Rosa and A. Piva, "Capacity of the watermark channel: how many bits can be hidden within a digital image", Proc. SPIE 3657, 1999, pp. 437-448.
- [3] P. Bas and J.M. Chassery, 'Tatouage d'image résistante aux transformations géométriques', 17ème colloque GRETSI, vannes, 13-17 septembre 1999.
- [4] M. Charrier, D. S. Cruz, and M. Larsson, "JPEG2000, the next millennium compression standard for still images," in Proc. IEEE Int. Conf. Multimedia Computing Systems'99, Florence, Italy, July 7-11, 1999, pp. 131-132.
- [5] I.J. Cox, J. Kiliani, T. Leighton and T. Shamon, 'Secure spread spectrum watermarking for multimedia', IEEE transaction on image processing June, (1997), pp. 1673-1687.
- [6] R. Dugad, K. Ratakonda, and N. Ahuja, 'A new wavelet-based scheme for watermarking images', Proceedings of the IEEE International Conference



- on Image Processing, ICIP '98, Chicago, IL, USA, October 1998.
- [7] B. G. Haskell, P. G. Howard, Y. A. LeCun, A. Puria, J. Ostermann, M. R. Civanlar, L. Rabiner, L. Bottou, and P. Haffner, "Image and video coding—Emerging standards and beyond," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 8, pp. 814–837, Nov. 1998.
 - [8] C.-T. Hsu and J.-L. Wu, "Multiresolution watermarking for digital images," *IEEE Trans. Circuits Syst.*, vol. 45, pp. 1097–1101, Aug. 1998.
 - [9] H. Inoue, A. Miyazaki, A. Yamamoto, and T. Katsura, "A digital watermark based on the wavelet transform and its robustness on image compression and transformation," *IEICE Trans. Fund. Electron., Commun., Comput. Sci.*, vol. E82-A, pp. 2–10, Jan. 1999.
 - [10] N. Kaewakamnerd and K.R. Rao, 'Wavelet based image adaptive scheme', *Electronics letters* Vol. 36, (2000), pp. 312-313.
 - [11] D. Kundur and D. Hatzinakos, "Digital watermarking using multiresolution wavelet decomposition," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, vol. 5, Seattle, WA, May 1998, pp. 2969–2972.
 - [12] D. Kundur and D. Hatzinakos, "A robust digital image watermarking method using wavelet-based fusion," in *Proc. 4th IEEE Int. Conf. Image Processing '97*, Santa Barbara, CA, Oct. 26–29, 1997, pp. 544–547.
 - [13] G. Nicchiotti and E. Ottaviani, "Non-invertible statistical wavelet watermarking," in *Proc. EUSIPCO'98*, vol. 4, Rhodes, Greece, Sept. 8–11, 1998, pp. 2289–2292.
 - [14] C.I. Podichuck and W. Zeng, 'image adaptive watermarking using visual models', *IEEE journal on selected area in communication*, Special Issue on Copyright and Privacy Protection Vol. 16, (1998), pp. 525-538.
 - [15] H. Seddik, E. B.Braiek, Color Medical Images Watermarking, *ICGST International Journal on Graphics, Vision and Image Processing* 6, Special Issue on Medical Image Processing (2006) 81-86.
 - [16] H. Seddik, M. Sayadi, F. Fnaiech and M. Cheriet, A New Spatial Watermarking Method, based on a Logarithmic transformation of An Encrypted embedded Mark, *Proc. IMACS Scientific Computation, Applied Mathematics and Simulation Conference*, , France, July 11-15, 2005.
 - [17] H. Seddik, M. Sayadi, F. Fnaiech et M. Cheriet, "Image Watermarking Based on the Hessenberg Transform" *Accepté dans: the International Journal of Image and Graphics*, Volume: 9, Issue: 3 (July 2009), pp 411-433.
 - [18] H-J. Wang, P-C. Su, and C.C. Jay Kuo, 'Wavelet-based digital image watermarking.' *Optics Express*, 3 pp. 497, December 1998.
 - [19] J. Wang and G. Wiederhold, "WaveMark: Digital image watermarking using daubechies' wavelets and error correcting codes," in *Proc. SPIE Int. Symp. Voice, Video, Data Communications*, Boston, MA, Nov. 1998.
 - [20] H.-J. M. Wang, P.-C. Su, and C.-C. J. Kuo, "Wavelet-based digital image watermarking," *Opt. Express*, vol. 3, no. 12, pp. 491–496, Dec. 7, 1998.
 - [21] R. B. Wolfgang, C. I. Podilchuk, and E. J. Delp, "Perceptual watermarks for digital images and video," *Proc. IEEE*, vol. 87, pp. 1108–1126, 1999.
 - [22] X.G. Xia, C. G. Boncelet, and G. R. Arce, 'Wavelet transform based watermark for digital images', *Optics Express*, 3 pp. 497, Dec. 1998.
 - [23] X. Xia, C. G Boncelet, and G. R. Arce, "A multiresolution watermark for digital images," *Proc. 4th IEEE Int. Conf. Image Processing '97*, CA, Oct. 26–29, 1997, pp. 548–551.
 - [24] K. Xiangui and J. Huang, "A DWT-DFT Composite Watermarking Scheme Robust to Both Affine ransform and JPEG Compression", *IEEE Transaction on circuits and systems for video technology* Vol. 13, no. 8, august 2003.
 - [25] W. Zhu, Z. Xiong, and Y.-Q. Zhang, "Mutliresolution watermarking for images and video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, pp. 545–550, June 1999.



Biography



Hassene seddik is born in 1970 in Tunisia (Tunis), he received the engineer degree from the national school from the national school of engineer since 1995, the master degree and the PhD degree in signal processing specially in image watermarking. He belongs

to the CEREP group in the high school of sciences and techniques of Tunisia. Dr. SEDDIK has published over 35 scholarly research papers in many journals and international conferences and he is already supervising over seven masters and 05 engineer ends of year projects. His research interests are focusing on two domains:

- Information security by applying cryptography and watermarking to voice, image and video.
- Voice, image and video processing, recognition, and filtering using neural networks applications.



Ezzedine Ben Braiek obtained his HDR on 2008 in Electrical engineering from ENSET Tunisia. He is, presently, professor in the department of electrical engineering at the technical university ESSTT and manager of the research group on vision and

image processing at the CEREP. His fields of interest include automatics, electronics, control, computer vision, image processing and its application in handwritten data recognition.







Modeling of the precipitations in radar and satellite images using the Autoregressive process

L. Sadouki, B. Haddad, D. Kemikem

*Laboratory of Image Processing and Radiation, Faculty of Computer Science and Electronics,
University of Sciences and Technology Houari Boumediene (USTHB),
PObox 32 El Alia, Bab Ezzouar, Algiers, Algeria
sdikleila@yahoo.fr, h_boualem@hotmail.com, kemikemds@yahoo.fr
http://www.usthb.dz*

Abstract

The aim of this paper is a modeling of precipitations in satellite and radar images that are useful for hydrologic applications. An investigation on a possible different description of precipitations over the sea and on the continent is attempted in images concerning the Mediterranean region. The model used in this work is the autoregressive process AR(p), where p is its order. This model makes possible the description of a series of 2208 satellite images, recorded during a period of 3 months via the IR03.9 Meteosat channel, and 4088 radar images recorded by the incoherent pulsed radar of sétif (Algeria). To conduct a detailed study of the phenomenon, a scanning area of 120×400 Km in the two areas (sea and continent) is adapted to examine all precipitation areas on the images, using a windows (5×5 pixels) process. The order of the AR model found in this paper, for the Mediterranean region and using the two types of images, is “1”, which is invariant in time and in space.

Keywords: Autoregressive, precipitations, Modeling, Images, Satellite, Radar.

1. Introduction

Over the past decade, great attention has been paid to the parameterization of the rainfall fields. In general, the most successful hydrological models are those that combine spatial and temporal variations of precipitations.[9]

Analyze a chronological series consists in finding an adequate mathematical model of the series evolution mechanism. The obtained model is used to fulfill the objectives such as prediction and control.

For a series which collect the satellite rainfall estimation, the methods published in literature are mainly based on the determination of precipitation rates associated with different types of clouds, or the characterization of atmospheric convection, or tracking the evolution of cloud [5] or evaluating the frequency of occurrence of

cold cloud [7]. However, satellite images do not ensure accurate identification of the cloud masses [1]. The weather radars are recognized as the most effective tool to detect precipitation fields in a given region. The radars images are widely used for hydrological purposes, rainfall structures and their danger identification for short-term forecasting use; also, they serve a rational and correct management of water resources.

Several statistical techniques allow to analyze rainfall data collected by networks of pluviometers were published in the literature [2, 3, 10]. Other researchers have shown that precipitation can be described by Markov chains of the first-order with two states taking into account the areas occupied in the radar images collected in Bordeaux (France) [8]. This representation is in the context of a global approach where we assumed that the observed state at the present moment depends only on the immediate previous state. In this paper, we propose a modeling of precipitations over the sea and on the continent of the Mediterranean region in satellite and radar images using the autoregressive process by taking into account additional previous states.

The remainder of the paper is organized as follows: Section (2) focuses on the mathematical formulation of an autoregressive process, sections (3) and (4) deal respectively with the database used in this work and the data processing. We illustrate and validate different results in sections (5) and (6). Finally, and in section (7), we give our conclusion.

2. Mathematical formulation

2.1. Autoregressive process of order p

Autoregressive models (AR), assume that X_t is a linear function of previous values.[6]

$$X_t = \sum_{i=1}^p \phi_i X_{t-i} + \varepsilon_t \quad (1)$$

Where ϕ_i , in this equation are the coefficients of auto-



regression.

The autocorrelation function is a very important function in the characterization of a linear process, it's written as follow [4]:

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \frac{\sum_{i=1}^N (X_i - \bar{X})(X_{i+h} - \bar{X})}{\sum_{i=1}^N (X_i - \bar{X})^2} \quad (2)$$

Where $\gamma(h)$ is the auto-covariance function and \bar{X} is the average.

This function provides information about the process memory, that is to say, the degree of dependence between observations at time t and those made at the time $t-h$.

2.2. The Autoregressive process coefficients

For $P = 1$, it is easy to see that $\phi_1 = \gamma(1)$ and $\sigma^2 = 1 - \phi_1^2$, for higher orders ($p \geq 2$), the coefficients of the AR process are calculated using the following recurrence relations : [4]

$$\phi_{hh} = \frac{\left[\gamma(h) - \sum_{p=1}^{h-1} \phi_{h-1,p} (h-p) \right]}{\sigma_{h-1}^2} \quad (3)$$

$$\phi_{h,p} = \phi_{h-1,p} - \phi_{hh} \phi_{h-1,h-p} \quad (4)$$

$$\sigma_h^2 = 1 - \sum_{i=1}^h \phi_{hi}^2 \quad (5)$$

If the process is of the order p , then:

$$\sigma_{p+1}^2 = \sigma_p^2$$

$$\phi_{p+1,i} = \phi_{p,i}$$

$$\phi_{p+1,p+1} = 0$$

The order of an autoregressive process is the value for which the partial autocorrelation function is zero. Estimated partial autocorrelation coefficients are considered invalid if they fall within the confidence interval constructed from the standard deviations. It is recognized that a series follow an AR (p) if the correlogram of its autocorrelation function decreases to 0 and its partial autocorrelation function vanishes beyond a lag p .

3. Database

To build our database, we considered three months of satellite and Radar observations in the Mediterranean region. The satellite images were taken by the channel IR03.9 of Meteosat (see Figure1.a), during the period of November, December 2009 and January for the 2010 year. For each day we recorded 24 images, the pixels of those images are coded over 8 bits and with a spatial resolution of 5km×5km.

The Radar images consist of 4088 images collected in Sétif in 1999. These images were taken by incoherent pulsed radar. (see Figure1.b). All images are recorded according to the PPI mode (Plan Position Indicator). They are represented in the format 512×512 pixels with a spatial resolution of 1km×1km. The number of color levels coding each pixel of these images is 16.

When we treated the satellite and radar images of the Mediterranean region we have seen that the echoes of rain on these images were characterized by high persistence. In addition, we found that the echoes of rain were standing over a period of more than a month.

Such properties imply that changes in precipitation over time can be described by a sequence of random variables forming an autoregressive process.

4. Data processing

In this study, the images are divided into two zones; one corresponds to the sea, the other one corresponds to the continent (see Figures 1.a and 1.b). For a detailed analysis of the phenomenon, we proceed to do a 5×5 pixel windows processing. To study a good part of our images, left and right shifts in both zones are applied.

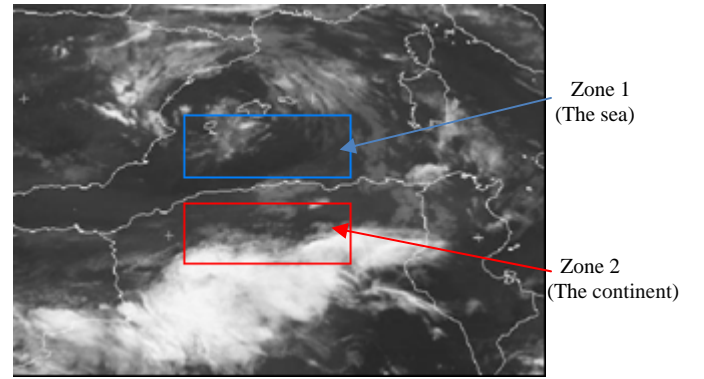


Figure 1.a. Satellite image of the Mediterranean area shows the two study zones

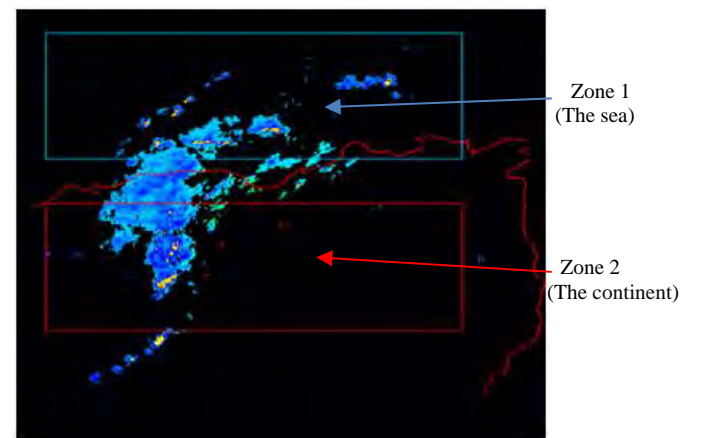


Figure 1.b. Radar image of the Mediterranean area shows the two study zones

The results versus time for both areas are represented by figures (2.a and 2.b). These curves are respectively obtained using the precipitations in the satellite images and the factors of reflectivity (ZdBZ) in the radar images.



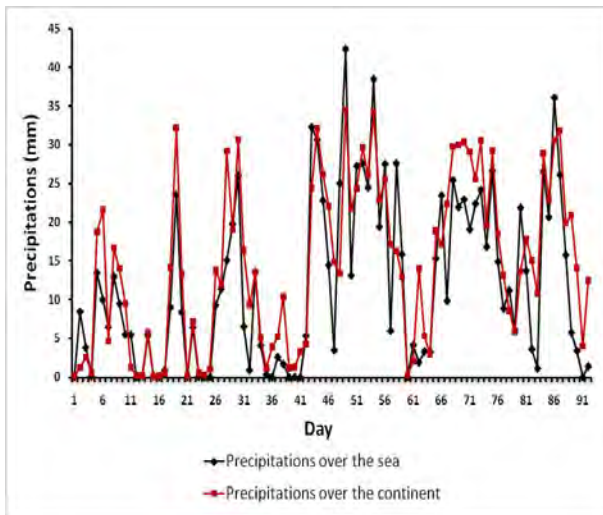


Figure 2.a. Daily precipitations over both sea (Zone 1) and continent (Zone 2) (Satellite images)

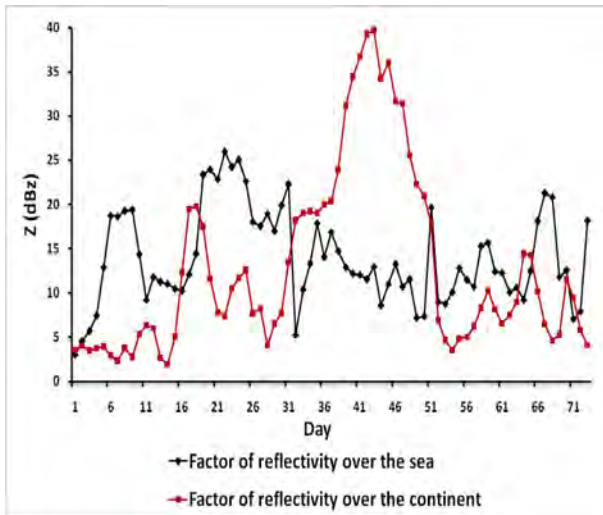


Figure 2.b. Daily factor of reflectivity over both sea (Zone 1) and continent (Zone 2) (Radar images)

From the previous curves, it is clear that the data series are stationary as they exhibit fluctuations in rainfall, around an average value, 13.27 for the precipitations given by the satellite images process and 13.44 for the reflectivity's factor given by the radar database.

5. Results and interpretations

The following figures show the variation of the first 15 autocorrelation coefficients (AC) and partial autocorrelation coefficients (PAC) for both areas.

The confidence band is given inside the interval $[-0.2, 0.2]$.

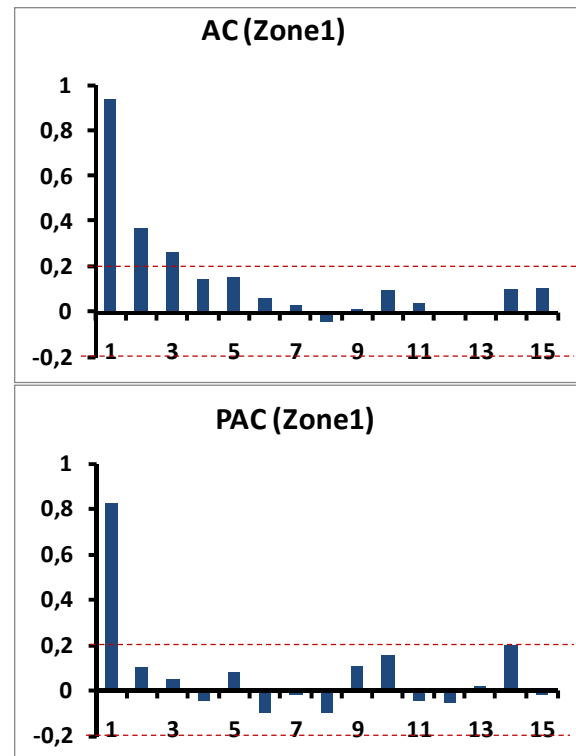


Figure 3.a. Variations of autocorrelation coefficients (AC) and partial autocorrelation (PAC).
(Process using satellite images for zone 1)

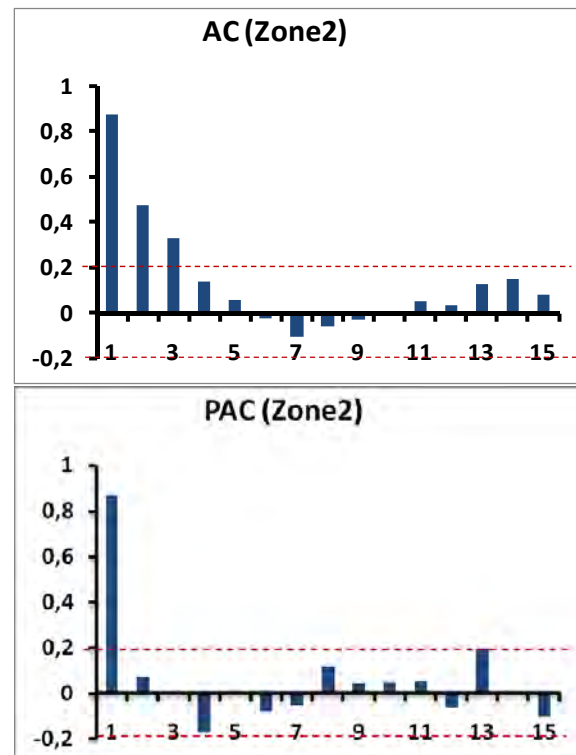


Figure 3.b. Variations of autocorrelation coefficients (AC) and partial autocorrelation (PAC).
(Process using satellite images for zone 2)



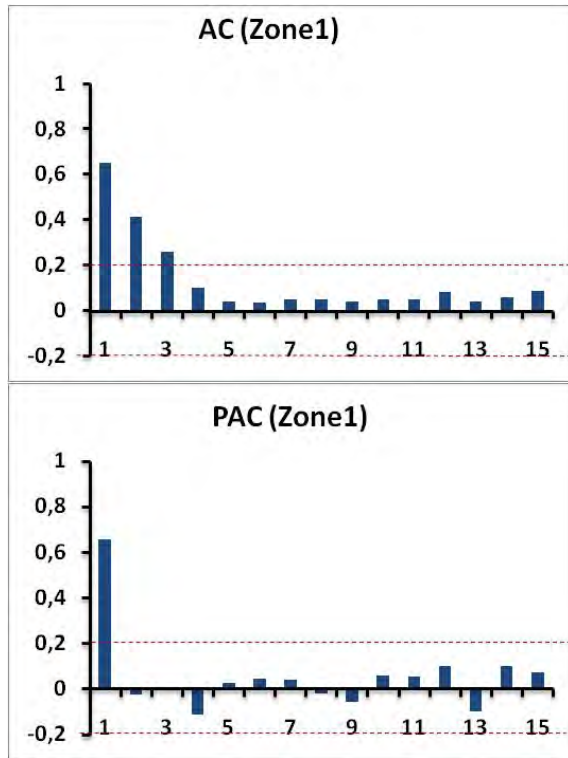


Figure 3.c. Variations of autocorrelation coefficients (AC) and partial autocorrelation (AC).
(Process using Radar images for zone 1)

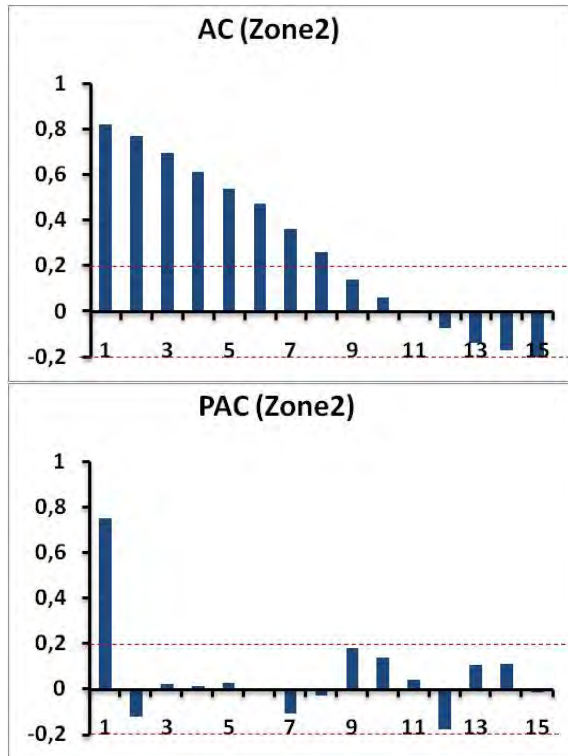


Figure 3.d. Variations of autocorrelation coefficients (AC) and partial autocorrelation (AC).
(Process using Radar images for zone 2)

We note that the autocorrelation coefficients decrease exponentially on the one hand and the partial

autocorrelation coefficients are within the confidence band defined by the interval $[-0.2, 0.2]$.

For the values $P=1, 2$ and 3 , the coefficients of the AR process are calculated and presented in the following table:

	Zone 1	Zone2
Processing by satellite images	AR(1): $\phi_{11} = 0.8235$ $\sigma_1^2 = 0.3218$	AR(1): $\phi_{11} = 0.8710$ $\sigma_1^2 = 0.2414$
	AR(2): $\phi_{21} = 0.7398$ $\phi_{22} = 0.1017$ $\sigma_2^2 = 0.4424$	AR(2): $\phi_{21} = 0.8099$ $\phi_{22} = 0.0701$ $\sigma_2^2 = 0.3391$
	AR(3): $\phi_{31} = 0.7350$ $\phi_{32} = 0.0674$ $\phi_{33} = 0.0464$ $\sigma_3^2 = 0.4530$	AR(3): $\phi_{31} = 0.8104$ $\phi_{32} = 0.0755$ $\phi_{33} = -0.0067$ $\sigma_3^2 = 0.3375$
Processing by Radar images	AR(1): $\phi_{11} = 0.6559$ $\sigma_1^2 = 0.5698$	AR(1): $\phi_{11} = 0.7500$ $\sigma_1^2 = 0.4375$
	AR(2): $\phi_{21} = 0.6729$ $\phi_{22} = -0.026$ $\sigma_2^2 = 0.5466$	AR(2): $\phi_{21} = 0.8430$ $\phi_{22} = -0.1240$ $\sigma_2^2 = 0.2740$
	AR(3): $\phi_{31} = 0.6731$ $\phi_{32} = -0.0314$ $\phi_{33} = 0.0081$ $\sigma_3^2 = 0.5459$	AR(3): $\phi_{31} = 0.8459$ $\phi_{32} = -0.1436$ $\phi_{33} = 0.0232$ $\sigma_3^2 = 0.2633$

Table 1. The coefficients of the AR process
(Satellite and Radar images).

The Autoregressive model was applied to all windows of the studied areas, where the parameter ϕ_{ij} varies between 0 and 0.871. The analysis of the partial autocorrelation coefficients ϕ_{ij} shows that:

- The process is stationary to the first order.
- The coefficients ϕ_{ip} of the zone continent are slightly higher than the coefficients of the zone sea.
- The variance of the error is larger over sea than over continent.
- The first order Autoregressive model (AR(1)) is well suited to describe precipitation.

6. Validation

After identifying the model, it is important to verify the adequacy of the model used with the observations. As the errors are not observable, they are replaced by the estimated errors calculated from the estimated parameters of model:



$$\hat{\varepsilon}_t = X_t - \sum_{i=1}^p \phi_{pi} X_{t-i} \quad (6)$$

The fit test will be based on the chi-square test and the autocorrelation function of the estimated errors. This function is written as:

$$\rho_h(\hat{\varepsilon}_t) = \frac{\sum_{t=1}^{n-h} \left(\hat{\varepsilon}_t \right) \left(\hat{\varepsilon}_{t+h} \right)}{\sum_{t=1}^n \hat{\varepsilon}_t^2} \quad (7)$$

However, because of some dispersion in the calculation of autocorrelation coefficients, it is better to consider them in their entirety. Thus, we calculate the first 15 autocorrelation errors.

Then we use the statistical distribution given by Box and Pierce [4]:

$$Q_n = n \sum_{h=1}^{15} \rho_h^2(\hat{\varepsilon}_t) \quad (8)$$

Where n is the total number of samples studied.

We compare the values of Q_n with the threshold χ^2 read on the table of chi-square, which is 23.68 for 14(15-P(=1)) degrees of freedom and a risk of error of $\alpha=0.05$.

Q_n	Zone 1	Zone 2
Process by Satellite data	20.53	18.32
Process by radar data	14.14	11.88

Table 2. Values of Q_n for both areas

The results given in Table 2 show that Q_n is still below the threshold χ^2 for both areas (sea and continent). The errors are independent of each other.

7. Conclusion

The aim of this paper is a modeling of the satellite and radar images precipitations that are useful for several applications and the study of the precipitations over the sea and on the continent of the Mediterranean region. The results of the statistical analysis of images showed that the rainfall is well described by a first order autoregressive process (AR(1)), irrespective of their types, over sea or land. This result is in good agreement with those published in the literature. This work showed that the historical rainfall data is limited only to the previous day. This work can be used in the agricultural field and in hydrology for the selection of dam sites.

8. Acknowledgements

The authors would like to thank all those who contributed to the satellite and radar data set used in this paper, notably Meteosat for the central-Europe and north-Africa images and the National Meteorology Office of Algeria for the Sétif data. We would like to thank also the reviewers for valuable comments and suggestions.

9. References

- [1] Ameer Z., Lazri M., Ameer S., Identification des masses nuageuses par fusion de données images (radar-satellite), Larhyss Journal, ISSN 1112-3680, n° 06, pp. 105-120, 2007.
- [2] Arnaud M., Contribution à l'étude stochastique markovienne des précipitations dans le bassin Adour-Garonne. PhD thesis, Toulouse (France), 1985.
- [3] Billingsley P., Statistical methods in Markov chains, In Stanford meetings of the Institute of Mathematical Statistics, Chicago (USA), 1960.
- [4] Bosq D et Lecoutre J.P., Analyse et prévision des séries chronologiques, Édition Chiron, Paris, 425 p, 1992.
- [5] Cadet D.L., Guillot B., Programme de recherche EPSAT (Estimation des Précipitations par satellite), Publication ORSTOM/Ministère de la coopération, 1991.
- [6] Didier D., Series temporelles– Modèles ARIMA, Séminaire EA "Sport – Performance – Santé", Mars 2000.
- [7] Guillot B., Satellite et précipitations : Contraintes techniques et physiques, analyse de quelques méthodes, problèmes de recherche et de validation, Veille Climatique Satellitaire, 55, 27-58, 1995.
- [8] Haddad B., Adane A., Mesnard F. and Sauvageot H., Modeling anomalous propagation using first-two stats Markov chain, Atmospheric research, 52, 237-244, 2000.
- [9] Haddad B., Sadouki L., Sauvageot H. et Adane A., Analyse de la dimension fractale des échos radar en Algérie, France et Sénégal. Télédétection, volum. 5, n°4, p. 299-306, 2006.
- [10] Mares I., A Markov chain for evaluation of moon rainfalls (in Romanian), studii si Cercetari de Meteorology, Part I/139, 203-219, 1974.



Biographies



Leila SADOUKI holds the engineer Diploma from the Computer Science and Electronics Faculty, University of Sciences and Technology of Algiers in 1997; she obtained the degree of Magister in 2003 in

Atmospheric radiation at the same university. From 1998 to 2003 she worked as Teacher Specialize in professional teaching. Since 2004 she is an assistant Professor at the Electrical and Electronic Engineering Institute of Boumerdes University. She's research interests are in radar and satellite images processing, neuronal networks, fuzzy logic and fractal approaches; at present, she is pursuing the PhD degree at the University of Sciences and Technology in the Image Processing and Radiation Laboratory.



Prof. Boualem HADDAD got his degree in Telecommunications Engineer from the national polytechnic school (Algeria) in 1982, his Magister degree in applied electronics in 1991 and his PhD degree in Atmospheric

radiation from the University of Sciences and Technology of Algiers in 2000. His areas of interest are electromagnetic radiation, weather radar and instrumentation and atmospheric modeling. He served as Chairman of the scientific council of the telecommunications Department from 2003-2004 and from 2007-2010. He is also, head and member of many research projects in Atmospheric Radiation, Electronics applied to meteorology and electromagnetic radiation in the low atmosphere. He published more than 14 research Journal papers in various fields, and more than 20 research papers in national and International Scientific Conferences. He has been teaching since 1984. From 1996-1999, he was the assistant director of education and pedagogy. He is currently a professor at the University of Sciences and Technology of Algiers since 2001, and member of Radiation and image processing laboratory.



Dallal KEMIKEM got her degree in control engineer in 1999 from the Faculty of Computer Science and Electronics, University of Sciences and Technology of Algiers. From 2003 to 2005, she was under training to be a

principal engineer in electronic instrumentation and exploitation of research reactors in the nuclear research center of Draria. From 2007 to 2010, she worked as principal engineer and main teacher of operating reactors for engineers and technicians staff in the same research center. She contributed in several international trainings in engineering reactors and other nuclear facilities since 2003 up to now. She got a degree of Magister in 2010 in Atmospheric radiation at the University of Sciences and Technology of Algiers. Currently, she is working as an Associate Researcher at the same center on nuclear instrumentation and control systems.





2D Segmentation of Intravascular Ultrasound Images: A Fast-Marching Method

Hassen Lazrag¹, Med Ali Hamdi^{1,2}, and Med Saber Naceur^{1,2}

¹Ecole Nationale d'Ingénieur de Tunis, Tunisie

²Institut National des Sciences Appliquées et de Technologie, Tunisie

mohamedalihamdi@yahoo.fr, hassen.lazrag@laposte.net, naceurs@yahoo.fr

<http://www.enit.rnu.tn.com>

Abstract

Intravascular ultrasound (IVUS) imaging is a catheter-based medical technique establishing itself as a useful tool for studying coronary atherosclerotic disease. It provides cross-sectional inside view of blood vessels that allows a complete quantitative assessment of their morphology, such as vascular wall, nature of atherosclerotic lesions or plaque shape and size. Manual processing of IVUS images is a tedious and time-consuming task and it is not efficient for clinical purposes. In this work, a new 2D IVUS segmentation model that incorporates a fast-marching algorithm and is based on gray level gradient of the vessel wall structures is developed. The fast-marching segmentation algorithm computes the lumen, intima plus plaque structure, and media borders simultaneously. Preliminary results of this new IVUS segmentation model agree very well with contours drawn manually by physicians. Moreover, fast-marching segmentation is appeared to be less sensitive to initialization.

Keywords: *Intravascular ultrasound (IVUS); Segmentation; Gray level gradient; Fast Marching; Coronary arteries*

1. Introduction

Coronary angiography is the gold-standard method for imaging and diagnosis of coronary atherosclerotic disease. However, it is restricted by its inability to quantify plaque burden beyond the luminal silhouette created by the angiographic contrast. In the last two decades, the intravascular ultrasound (IVUS) imaging has been proven better in the assessment of coronary atherosclerosis [1].

Intravascular ultrasound (IVUS) technology is a relatively new imaging technique that can offer information complementary to angiography technology, while IVUS provides an inside view of the arterial wall; coronary angiography illustrates only the silhouette of that vessel in the body [2, 3].

IVUS is a catheter-based technique that provides two-dimensional (2D) cross-sectional images of coronary

artery and, therefore, accurate information concerning the lumen, wall and plaque area, and vessel morphology. Most importantly, the tomographic nature of intravascular ultrasound makes 3D reconstruction of the vessel wall possible. For this purpose, fusion between biplane angiography and IVUS to recover the catheter path provides geometrically accurate 3D reconstruction. However, a typical intravascular ultrasound bull-back produces several hundred of images making non-automatic analysis of the data long, fastidious and subject to intra- and interobserver variabilities. These could be serious limitations against the clinical usage of IVUS technique. Therefore, because of poor IVUS image quality due to the existence of speckle noise, imaging artifacts, calcifications shadowing and rupture of parts of the vessel wall, it is necessary to develop automatic segmentation methods taking into account the nature of IVUS bull-backs.

Several approaches for semi- and full automatic IVUS segmentation have been proposed so far, including texture analysis [4], active contours [5], knowledge-based graph searching [6], minimum cost algorithms [7], and region growing [8]. Among them, the active contour model and its variations showed remarkable feasibility and accuracy [9–11]. However, some of these algorithms require manual contour initialization, thereby increasing the user interaction, and concomitantly the uncertainty of the segmentation. In the present study, usage of a local image characteristic, i.e. the gray level gradients of the IVUS image, combined to a fast-marching segmentation model, a local-based method handling topological contour changes is proposed to address the IVUS segmentation problem.

2. Methods

Image preprocessing: The preprocessing [16] of the image data consists of three steps: (a) the determination of the region of interest (ROI), (b) representation of the images in polar coordinates, and (c) reduction of noise.



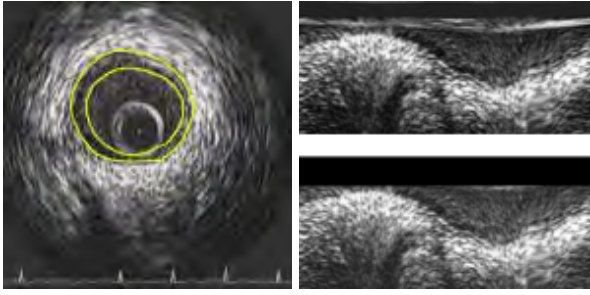


Figure 1: Original IVUS image (left) and corresponding polar coordinate images before (right top) and after (right bottom) the removal of catheter-induced artifacts.

The first step consisted in determining the image zones where the regions of interest (ROI) are extracted yielding a squared image.

The transformation of the IVUS images in polar coordinates is important to facilitate efficient pre-processing of local image regions in terms of their radial and tangential directions. Besides, it facilitates a number of other pos-processing steps, such as contour initialization and the detected contour smoothing. Consequently, each of the original IVUS images is transformed to a polar coordinate image where columns and rows correspond to angle and distance from the center of the catheter, respectively. The polar coordinate image, denoted $I(r, \theta)$, is used for the remaining analysis process.

The images produced by IVUS bull-backs include not only tissue and vessel regions but also the transducer of the catheter itself. The latter defines a dead zone of radius equal to that of the transducer, where no useful information is contained. Because its good results for speckle reduction and edge enhancement, an adaptive median filter of size 5X5 is applied to remove these catheter-included artifacts.

Fast-Marching Segmentation Model:

Fast-Marching Methods and Level Sets Methods are computational techniques for tracking propagating interfaces. They share the virtues of working in an arbitrary number of spaces dimensions with no change, handle topological merger and splitting with no special procedures, and accurately and efficiently compute the motion of fronts moving under complex speed laws, including large variations in velocities and shape discontinuities.

Fast Marching Methods, introduced by Sethian in [14], approximate the solution of a boundary value partial differential equations view of propagating interface, while level set methods, introduced by Osher and Sethian in [12], approximate the solution of initial value partial differential equation. At the core, both techniques rely on viscosity solutions for Hamilton-Jacobi equations, upwind schemes for hyperbolic conservation laws, and theory of curves and surface evolution developed by Sethian in [12].

Consider a boundary, either a curve into two dimensions or a surface in three dimensions, separating on region from another. The speed function F , which used to advance this interface may depend on a wide variety of factors, can be written as:

$$F = F(L, G, I) \quad (1)$$

- Local image properties: they are determined by local geometric information, such as curvature and gradient gray level.
- Global image properties: they depend on the shape and position of the interface front.
- Independent properties: they are those independent of the shape of the front.

The difficult challenge in interface problems is to guarantee that the produced speed function model F is adequate.

Level Set Methods: Level Set Methods view a moving front as the zero level-set of a function $\phi(x, y, t)$. The evolution equation for the interface moving with speed F in its normal direction can be written [13] in the following general form

$$\frac{\partial \phi(x, y, t)}{\partial t} + F|\nabla \phi| = 0 \quad (2)$$

The level-set model can be applied to image segmentation by interpreting image contours as the propagating interface final position [13, 14]. To achieve this, the speed function should become close to zero when the propagating front meets with the desired contour thus making the interface stop at this position and then ending the segmentation process.

The key idea in level set methods is to approximate the gradient in the level set formulation in a way that constructs the correct entropy-satisfying condition. One of such schemes is given in [14], namely

$$\phi_{i,j}^{n+1} = \phi_{i,j}^n - \Delta t F_{i,j} \left(\frac{\max(D_{i,j}^{-x} \phi, 0)^2 + \min(D_{i,j}^{+x} \phi, 0)^2}{\max(D_{i,j}^{-y} \phi, 0)^2 + \min(D_{i,j}^{+y} \phi, 0)^2} \right)^{1/2} \quad (3)$$

Fast Marching Methods: Fast Marching Method can be described as the evolution of an interface propagating in the same direction with a speed function F that is neither strictly positive nor negative. In this case, the evolving boundary must be inside the region to segment for a positive speed (or outside the desired region for a negative one). The central idea behind the fast-marching formulation is to systematically construct the arrival time $T(x, y)$, which is the time when the evolving contour crosses the point (x, y) . The T function satisfies the partial differential equation (4), stating that the arrival time difference between two neighbour pixels increases as the speed of the contour decreases.

$$|\nabla T| F = 1 \quad (4)$$

As discussed earlier, the fast marching requires careful construction of upwind, entropy-satisfying scheme to the gradient approximation. Using the above scheme, we have



$$\left[\frac{\max(D_{i,j}^{-x}T,0)^2 + \min(D_{i,j}^{+x}T,0)^2}{\max(D_{i,j}^{-y}T,0)^2 + \min(D_{i,j}^{+y}T,0)^2} \right]^{1/2} F_{i,j} = 1, \quad (5)$$

The propagation of the front is done via the construction of the arrival time function T . The construction algorithm [14] selects the front point having the smallest value of T and calculates the arrival times of its four neighbours, and so on until the front has propagated across the entire image. The front is considered stationary when the arrival time gradient is sufficiently high.

Since multiple borders (lumen-intima and media-adventitia) must be detected on the IVUS images, the fast-marching segmentation algorithm is done via a multiple interface extension [15]. Every interface associated to the vessel wall contours evolves at a velocity defined in terms of gray level gradient image feature. The speed function is given by Sethian [13]:

$$F(i, j) = \frac{1}{1 + |\nabla G_\sigma * I_s|} \quad (6)$$

where G_σ is a 9x9 pixel symmetric Gaussian smoothing filter of standard deviation $\sigma = 0.5$ and the gradient is computed in 2D. Therefore, this F function speeds up the interface fronts on low gradient regions.

Segmentation initialization: The fast marching segmentation method requires an initialization procedure which is necessary to detect the pixels which are likely to belong to the lumen-intima and media-adventitia borders. Two initialization steps were defined to this end, each used for the initialization of the respective border.

At the first initialization step, the fast marching based method is applied to IVUS image to detect intima border by defining a circle of radius 1.5 times larger than radius of the transducer located at the centre of the (ROI) image. The initial contour can be located easily due to the catheter clearly different contrast with the surrounding region.

The speed F in equation (6) should take negative value in order to speed up the expansion of the curve after some iteration to locate the intima boundary. In the second step, a large circle is placed outside the IVUS image and with a positive value for F the curve can shrink faster to the media-adventitia border after some iteration.

3. Results

Segmentation results for the gray level gradient based fast marching method are shown in Figure 2 and Figure 3. The lumen and media detected boundaries are presented for typical cross-sectional IVUS images of the whole in-vivo and simulated pullbacks.

The results are promising and they demonstrate the fast marching accuracy in determining lumen and plaque contours. The media-adventitia border detection depends extremely on the quality of the images so a more accurate set of results depends directly on the strategy followed the contour initialization and detection.

The fast marching segmentation model is applied first to simulated IVUS images and then to real IVUS images.

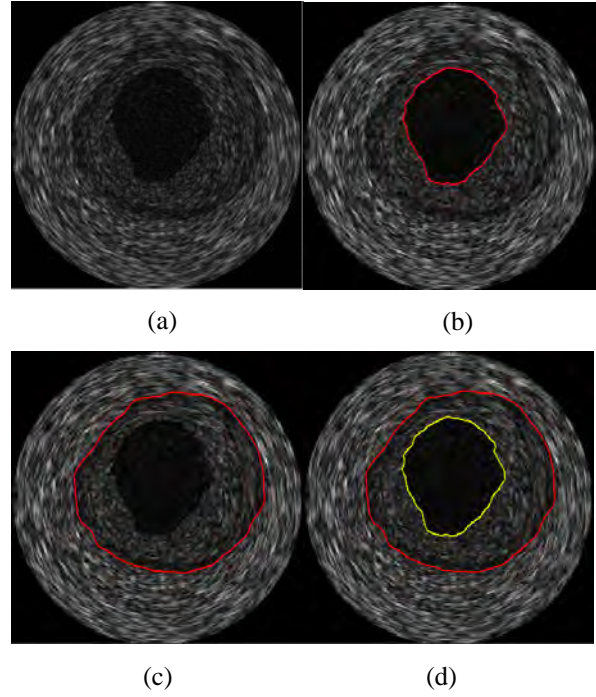


Figure 2: Simulated IVUS (a) Intima border detection (b) media-adventitia border detection (c) Intima-LEE border detection (d).

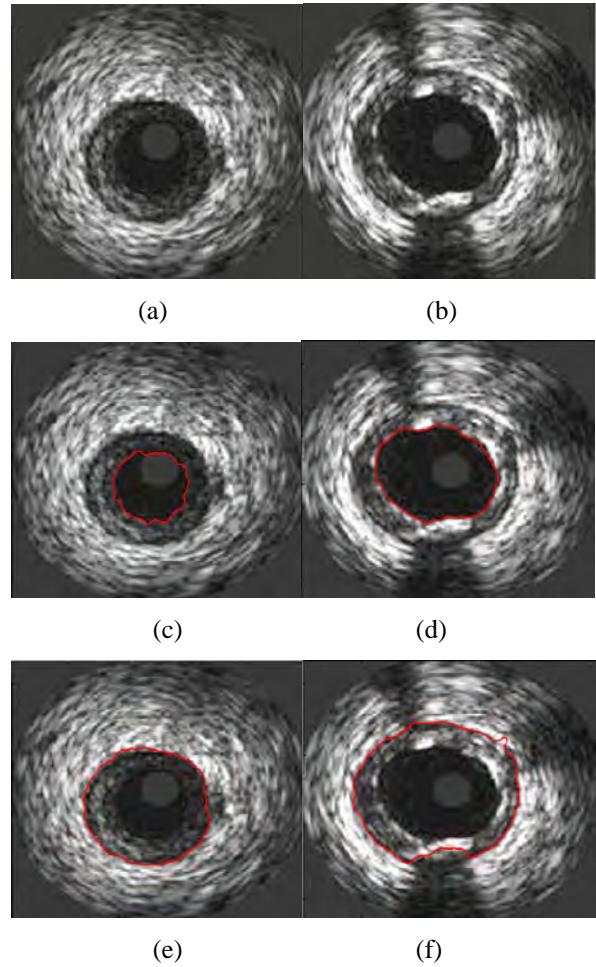


Figure 3: Real IVUS images (a),(b) Intima border detection (c),(d) media-adventitia border detection (e), (f).



Results show the accuracy of our methods in detecting lumen border and media-adventitia contour. As shown in Fig. 2, the results of this method for a simulated IVUS image of coronary artery, the boundaries of the vessel wall were successfully extracted. Fig. 3 shows an example outcome of the method, the intima and medial-adventitial borders were detected despite the presence of noise.

4. Conclusion

The goal of this work is to demonstrate the IVUS segmentation potential of the fast marching method, and the usefulness of local image features such as gray level gradients of vessel wall anatomical structures. Fig. 3 showed that vessel wall boundaries can be identified even when image contrast is not very high and when the luminal contour shape is irregular.

Qualitative analysis of the detected contours indicates that the level set is an accurate segmentation method for intravascular ultrasound imaging, but this should be further confirmed from validation with several manually traced contours by independent experts.

These preliminary segmentation results showed that this is a promising technique for IVUS image processing. Further investigation will be focused on improving the algorithms speed and on new methods to automate the initialization step to minimize user's interactions.

5. References

- [1] G.S. Mintz, S.E. Nissen, W.D. Anderson, S.R. Bailey, R. Erbel, P.J. Fitzgerald, F.J. Pinto, K. Rosenfield, R.J. Siegel, E.M. Tuzcu, P.G. Yock, American College of Cardiology clinical expert consensus document on standards for acquisition, measurement and reporting of intravascular ultrasound studies (IVUS), *J. Am. Coll. Cardiol.* 37 (2001) 1478–1492.
- [2] Rotger D, Radeva P, Canero C, Villanueva JJ. Corresponding IVUS and angiogram image data. In: *IEEE Proc Comput Cardiol.* 2001.
- [3] Wahle A, Prause GPM, DeJong SC, Sonka M. Geometrically correct 3-D reconstruction of intravascular ultrasound images by fusion with biplane angiography-methods and validation. *IEEE Trans Med Imaging* (1999);18-19.
- [4] A. Mojsilovic, M. Popovic, N. Amodaj, R. Babic, M. Ostojic, Automatic segmentation of intravascular ultrasound images: a texture-based approach, *Ann. Biomed. Eng.* 25 (1997) 1059-1071.
- [5] R. Cothren, R. Shekhar, M. Tuzcu, S. Nissen, F. Cornhill, G. Vince, Three-dimensional reconstruction of the coronary artery wall by image fusion of intravascular ultrasound and bi-plane angiography, *Int. J. Card. Imaging* 16 (2000) 69–85.
- [6] M. Sonka, X. Zhang, M. Siebes, M. Bissing, S. DeJong, S. Collins, C. McKay, Segmentation of intravascular ultrasound images: a knowledge based approach, *IEEE Trans. Med. Imaging* 14 (1995) 719–732.
- [7] G. Kovalski, R. Beyar, R. Shofti, H. Azhari, Three-dimensional automatic quantitative analysis of intravascular ultrasound images, *Ultrasound Med. Biol.* 26 (2000) 527–537.
- [8] E. Brusseau, C.L. de Korte, F. Mastik, J. Schaar, A.F. van der Steen, Fully automatic luminal contour segmentation in intracoronary ultrasound imaging-A statistical approach, *IEEE Trans. Med. Imaging* 23 (2004) 554–566.
- [9] J.D. Klingensmith, R. Shekhar, D.G. Vince, Evaluation of three dimensional segmentation algorithms for the identification of luminal and medial-adventitial borders in intravascular ultrasound images, *IEEE Trans. Med. Imaging* 19 (2000) 996–1011.
- [10] Giannoglou, G.D., Chatzizisis, Y.S., Koutkias, V., Kompatsiaris, I., Papadogiorgaki, M., Mezaris, V., Parissi, E., Diamantopoulos, P., Strintzis, M.G., Maglaveras, N., Parcharidis, G.E., Louridas, G.E.: A novel active contour model for fully automated segmentation of intravascular ultrasound images: In vivo validation in human coronary arteries. *Comput. Biol. Med.* 37 (2007) 1292–1302
- [11] Sanz-Requena, R., Moratal, D., Garcia-Sanchez, D.R., Bodí, V., Rieta, J.J., Sanchis, J.M.: Automatic segmentation and 3d reconstruction of intravascular ultrasound images for a fast preliminar evaluation of vessel pathologies. *Comput. Med. Imag. Graph.* 31 (2007) 71–80.
- [12] Malladi, R., Sethian, J.A. Vemuri, B.C.: Shape Modeling with Front Propagation: A Level Set Approach. *IEEE T Pattern Anal.* 17 (2), pages 158-175, 1995.
- [13] Sethian, J.A.: Level Set and Fast Marching Methods. *Evolving Interfaces in Computational Geometry, Fluids Mechanics, Computer Vision and Materials Science.* 2ed. Cambridge, UK, Cambridge University press, 1999.
- [14] Sethian, J.A.: A fast marching level set method for monotonically advancing fronts. *Proceedings of the National Academy of Sciences of the United States of America*, vol. 93, pages 1591-1595, 1996.
- [15] Sifakis, E., Garcia, C., Tziritis, G.: Bayesian Level Sets for Image Segmentation. *J Vis Commun Image R*, vol. 13, pages 44-64, 2002.
- [16] Hamdi M.A. :Combining an alternating sequential filter (ASF) and curvelet for denoising coronal MRI image. *Contemporary engineering sciences.* Vol.5, 2012 ,n02, 85-90





Hassen LAZRAG was born in 1975. He received the Diploma degree in Electrical and Computer Engineering and the M.Sc. degree in Signal and Image Processing from the

National Institute of Engineering of Tunis, Tunisia, in 2001, and 2002, respectively. Since 2008, he is a Ph.D. candidate and a Research Assistant with LTSIRS Lab, Tunis, Tunisia. His research interests include medical image analysis /processing.



Mohamed Ali HAMDI received the Ph.D. degree in image processing, from the National University of engineering of TUNIS; He is Assistant Professor of Applied and Computational Mathematics and

electronic, National institute of applied sciences and technology. His research interests are in the areas of digital signal processing (DSP), statistical estimation and their applications to signal and image processing and scientific computing.

Mohamed Saber NACEUR is an associate Professor and head of physics and instrumentation department with National institute of applied sciences and technology. He Co-directs the laboratory of remote sensing data and information systems to space reference (LTSIRS) wich includes forty researchers. His research relates to the treatment and analyzes multimode images of remote sensing data and the geographical information systems (GIS).







Using three-dimensional (3D) Sierpinski Gasket to Generate and Recursively Re-Generate 3D Fractals - Closing the Self-Similarity Loop

Bulusu Rama and Jibitesh Mishra

Department of Computer Science & Engineering, MLR Institute of Technology, Hyderabad, INDIA
Department of Information Technology, College of Engineering & Technology, BPUT, Bhubaneswar, INDIA
 rama_bulusu@yahoo.com, mishrajibitesh@gmail.com

Abstract

Fractals provide an innovative method for generating three-dimensional (3D) images of real-world objects by using computational modelling algorithms based on the imperatives of self-similarity, scale invariance, and dimensionality. Images such as coastlines, terrains, cloud mountains, and most interestingly, random shapes composed of curves, sets of curves, etc. present a multi-varied spectrum of fractals usage in domains ranging from multi-coloured, multi-patterned fractal landscapes of natural geographic entities, image compression to even modelling of molecular ecosystems. Fractal geometry provides a basis for modelling the infinite detail found in nature. Fractals contain their scale down, rotate and skew replicas embedded in them. Of the many different types of fractals that have come into limelight since their origin, the Sierpinski fractal has eluded both mathematicians and computer scientists alike. And the two-dimensional (2D), 3D, etc. versions of the same have been realized based on the starting axioms/generators as either triangle/pyramid or square/cube. The resulting fractals are Sierpinski Triangle and Sierpinski Pyramid in case of the triangle-based generations; and Sierpinski Carpet and Sierpinski Gasket in case of the square/cube based fractals. This paper describes a methodology that illustrates the closeness in self-similarity otherwise termed as closing the self-similarity loop - of fractal images by way of generating a 3D Sierpinski Gasket fractal starting with a cube as the baseline shape and applying the same algorithm recursively by way of Iterated Function System (IFS)-like transformations of ((x,y) rotation, z (zoom)) - changing a newly introduced property called depth, from 3 to 2 to 1 in that order - to arrive at a new 3D Sierpinski Gasket that resembles a self-similar copy of the original 3D cube-based Sierpinski Gasket [9],[10],[11]. The depth property represents an additional variant which represents an iterative- recursive execution count.

Keywords: *Fractals, 3D Images, Sierpinski Traingle, Sierpinski Pyramid, Sierpinski Carpet, Sierpinski Gasket, 3D rendering, Recursion, IFS.*

Nomenclature:

I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism - *Fractals, Visible Line/Surface Algorithms*; I.3.3 [Computer Graphics]: Picture/Image Generation - *Display Algorithms, Viewing Algorithms*; G.1.2 [Numerical Analysis]: Approximation - *Wavelets and Fractals, Special Approximation Functions*

1. Introduction

A fractal is a rough or fragmented geometric shape that can be subdivided into parts, each of which is (at least approximately) a reduced size copy of the whole or in other words, is self-similar when compared with respect to the original shape. The term was coined by Benoit Mandelbrot in 1975 and was derived from the Latin word "fractus" meaning "broken" or "fractional". The primary characteristic properties of fractals are self-similarity, scale invariance and general irregularity in shape due to which they tend to have a significant detail even after magnification-the more the magnification the more the detail. In most cases, a fractal can be generated by a repeating pattern constructed by a recursive or iterative process. Natural fractals possess statistical self-similarity whereas regular fractals such as Sierpinski Triangle, Sierpinski Gasket, Cantor set or Koch curve contain exact self-similarity. The 3D renderings of Sierpinski Fractal with Triangle as the initial axiom/generator are the Sierpinski Traingle and set of n-dimensional Sierpinski Pyramid(s) where $n \geq 3$. Analogously, the 3D renderings of the Sierpinski Fractal with Square as the starting axiom/generator are the Sierpinski Carpet and the set of n-dimensional Sierpinski Gaskets, where $n \geq 3$. Paul Bourke has studied the Sierpinski Gasket and its versions in [17]. In this paper the authors describe a methodology that attempts to illustrate the concept of closing the self-similarity loop of fractal images by way of generating a 3D Sierpinski Gasket fractal starting with a cube as the baseline shape and applying the same algorithm recursively by way of IFS-like transformations of ((x,y) rotation, z (zoom)) - changing the recursion depth from 3 to 2 to 1 in that order. The source image for



the 3D Gasket is generated based on the 2D version of the same that is obtained by applying IFS transformations to a square figure after n iterations where $n \geq 3$. The original C program - courtesy of Andrey Mirtchovski (aam396@mail.usask.ca) - is used in the above implementation [15]. The displayed output of the same based on a typical set of inputs is resented. The remainder of the paper is organized as follows: Section (2) focuses on brief description of Sierpinski Triangle, Sierpinski Pyramid, Sierpinski Carpet and Sierpinski Gasket. Section (3) emphasizes on method of recursively generating the 3D Sierpinski Gasket to demonstrate closing the self-similarity loop. Section (4) shows the experimental results and Section (5) provides the concluding remarks.

2. About Sierpinski Triangle, Sierpinski Pyramid, Sierpinski Carpet and Sierpinski Gasket

A brief description of Sierpinski Triangle, Sierpinski Pyramid, Sierpinski Carpet and Sierpinski Gasket along with their properties followed by a description of generating the 2D and 3D versions of the same by using iterative function systems is presented in the paragraphs that follow. A detailed account of the Sierpinski Triangle and Sierpinski Gasket can be found in the book(s) as in [14] and in [9]-[10].

Sierpinski Gasket/Pyramid

The Sierpinski gasket, also known as Sierpinski triangle is a fractal produced by starting with one big filled-in triangle and dividing the triangle into four smaller triangles, then cutting out smaller and smaller ones from it in appropriate places in a recursion process which involves blanking out the center triangle and carrying out the above recursion process to operate upon the peripheral triangles. This involves a starting figure, a generator (the transformation) and a recursive process of applying the generator to the starting figure. The Sierpinski Gasket can also be generated by starting with a rectangle as the starting figure and applying a set of affine transformations (a rotation, a scaling and a translation) in each iteration step. The resulting fractal is the Sierpinski Gasket, a self-similar, equilaterally shaped triangular image that is symmetric with respect to the starting triangle.

The **Sierpinski Pyramid** is a three dimensional version of the two dimensional Sierpinski Gasket. Like the Sierpinski gasket, it is a recursive structure. A recursive structure of power 3 generates a Sierpinski Pyramid in 3D. Essentially this means that the 3D Sierpinski Pyramid consists of a pyramid constructed out of 4, power 2 Sierpinski Gaskets.

Sierpinski Carpet

The Sierpinski carpet is a fractal produced by removing the central square piece from a square sliced into thirds, horizontally and vertically. The area tends to zero and the total perimeter of the holes tends to infinity. Recursively cutting out smaller and smaller ones from it in the appropriate places yields the Sierpinski Carpet.

3. Method of Recursively Generating the 3D Sierpinski Gasket to Demonstrate Closing the Self-similarity Loop

A step-by-step description of the method used to generate and re-generate 3D fractals based on the cubed-based version of the 3D Sierpinski Gasket is outlined in the following steps:

As Figures 1, 2, 3 suggest, the originally generated 2D Sierpinski Gasket (Figure 1.3) uses a square as initiator (Figure 1.1) and the IFS as generator (Figure 1.2). The equivalent 3D version of the Sierpinski Gasket (Figure 2.2.1) uses a cube as initiator (Figure 2.1.1) and the IFS as the generator (Figure 2.1.2) all of them having an additional property called depth = 3; Figures 2.2.2, 2.3 and Figure 3 are the maneuvered versions of Figure 2.2.1 obtained by recursive re-generation with the same depth property (i.e., depth=3). Figure 4 is the maneuvered version of Figure 3 obtained by recursive re-generation with the depth property equal to 2. Finally Figure 5 seemingly displays the starting 3D Sierpinski Gasket as a 3D Cubed version of Figure 2 with the depth property changed to 1. The step-by-step procedure is as follows:

1. Adding a new depth property that is input-dependent to a chosen 3D Cube image. The default depth is chosen to be 3.
2. The generating IFS consists of a (x,y,z) transformation that consists of an (x, y) rotation and a z -based zoom or the scaling factor.
3. Setting the reference frame by specifying a (height, width) pair to the enclosing (movable) window. The default colour scheme is red/black for each cube-based 3D fractal.
4. Using the IFS, auto-generate 8 self-similar fractal blocks and position them at the 8 different corners of its parent cube fractal. This creates a linked chain of smaller but self-similar 3D Sierpinski Gaskets – that are within the boundaries of the frame-of-reference.

The dynamic variables involved in this method are:

1. The depth property that **represents an additional new variant that defines a projection**
2. The reference frame itself
3. The (x,y,z) triple representing the translation for the 3D fractal. Here z represents the zoom i.e., scaling factor

The dynamic translation is achieved by auto-capturing the variations in the scaling (the z -variable) value and the corresponding (x,y) rotation; and auto-adapting the GL projection(s) to the conforming viewports.

The program runs as a Windows console application using the Open Glut API, with the 3D Sierpinski Gasket being generated as an IFS-based recursive version of the 2D Sierpinski Carpet that uses a cube as the base initiator for the IFS. Successive fractal images are obtained by interactive mouse-based positional translation and context-aware zoom-based scaling in a dynamic fashion and varying the depth property from 3 to 1. Using MATHLAB 3D Image Rendering software, the displayed graphics are processed for 'true' GUI compatibility.

The original 3D Sierpinski Gasket (Cube based version) is shown in Figure 2.2.1. This has a depth of 3. By



mouse-over shearing and skewing of the same, the algorithm re-generates Figures 2.2.2, 2.3, 3 all of which also simulate a depth of 3. Figure 4 shows the 3D Sierpinski Gasket obtained by applying the above angular shearing to emulate a depth of 2. The final Figure 5 shows the 3D Sierpinski Gasket obtained by applying the above angular shearing to emulate a depth of 1. This proves that the 3D Sierpinski Gasket of depth 3 and depth 1 are almost identical in geometric similarity – a result that demonstrates the closeness in self-similarity of fractals. In this case, the self-similarity was symmetric in nature.

The IFS used for the experimental results is presented in the details that follow.

1. First we get the 2D Sierpinski Gasket followed by GL DEPTH BUFFERING and a set of recursively iterated IFS to get to the 3D Cube.
2. The IFS Code of the 3D Sierpinski Gasket (as discussed in this paper) consists of an Initiator (or axiom) and a generator (i.e., a set of rules) that are applied as an IFS and recursively executed n times where $1 \leq n \leq 9$. The degree of iteration or recursion is termed as depth of this variable.

The experimental results demonstrated here use depths 3, 2, 1 respectively. Also the output of depth n based 3d Sierpinski Gasket is used as the input for the depth $n-1$ based 3D Sierpinski Gasket.

//Initiator: Cube with the following definition:

Recursive depth = 3;

Cube Side = 10;

(x,y,z) = (0,0,0); // Translation

// Draw a cube of size= n . Each triple corresponds to a
// distinct edge of a cube.

float boundary[8][3] = {

{-1, -1, -1},

{-1, 1, -1},

{ 1, -1, -1},

{ 1, 1, -1},

{-1, -1, 1},

{-1, 1, 1},

{ 1, -1, 1},

{ 1, 1, 1},

};

Draw_CUBE()

{

// the pseudo code goes here.

}

//Generator: the function with following set of rules:

Draw_RecursiveDepthCUBE(int depth) is the generator. Here depth is recursive depth (i.e., the number of iterations deep the generator can be applied to itself as a recursive call to generate different versions of the 3D Sierpinski Gasket without losing its self-similarity.

//The zoom factor or the z-buffer (depth buffer

//corresponding to the GL_DEPTH BUFFERING IS

// TREATED SEPARATELY FROM THIS 'DEPTH'

//VARIABLE

```
{
  Axiom = Draw_CUBE();

  F = Draw_RecursiveDepthCUBE(1<=depth<=9);

  F3 = Draw_RecursiveDepthF3CUBE(3);

  F2 = Draw_RecursiveDepthF3CUBE(2);

  F1 = Draw_RecursiveDepthF2CUBE(1);

} RECURSIVELY_RE-GENERATE_3DSIERPINSKI;
```

Additional details and other related techniques for further research can be found in the book(s) as in [14] and in [7]-[10].

4. Experimental Results

This section depicts the results obtained by applying above described method(s) to demonstrate one scenario of closing the fractal self-similarity loop. One conclusion that can be made in a definitive manner from this experiment is: We started with a cube to generate the 3D Sierpinski Gasket (by imitating the 3D Sierpinski Triangle) and then recursively re-generated the 3D fractal to arrive a new 3D Sierpinski Gasket version - that resembles an almost 360-degree self-similar 3D version of the original fractal - but one that is constructed as a self-manipulated syndicate of successive iterated version-results, each of which represents a 3D fractal-pore of the initial 3D Sierpinski Gasket. The program was written in the C++ programming language, the 3D “visual” rendering was done using the Open Glut Graphics library. It runs as a Windows console application with the variable depth being input as the first command-line argument (any no-zero integer value, preferably from 1 to 9).

Figures 1.1 to 1.3 show iterated version displays of generating the 2D Sierpinski Gasket after 0, 3, n iterations. Figure 2.2.1 shows the 3D Sierpinski Gasket with depth 3 (the default) generated using a cube as the initializing axiom (Figure 2.1.1) and the IFS-based generator as shown in Figure 2.1.2. Figures 2.2.2 and Figure 3 show the recursively re-generated versions of the same IFS. Figure 3 is used as primary input source for the demonstration process. Figures 4 and 5 depict the 3D versions of the Sierpinski Gasket with depths 2 and 1 respectively.



Figure 1.1 Initial 2D Sierpinski Gasket – Square.



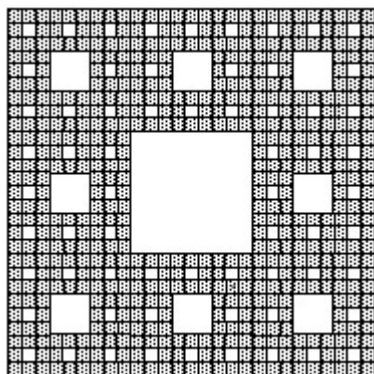


Figure 1.2 2D Sierpinski Gasket after 3 IFS-iterations applied to Figure 1.1.

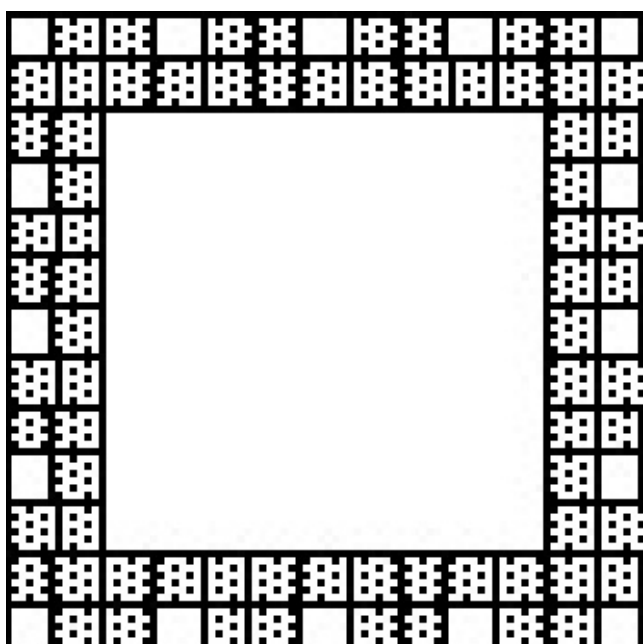


Figure 1.3 2D Sierpinski Gasket, after applying IFS transformations randomly (to Zoom-in) on Figure 1.2.

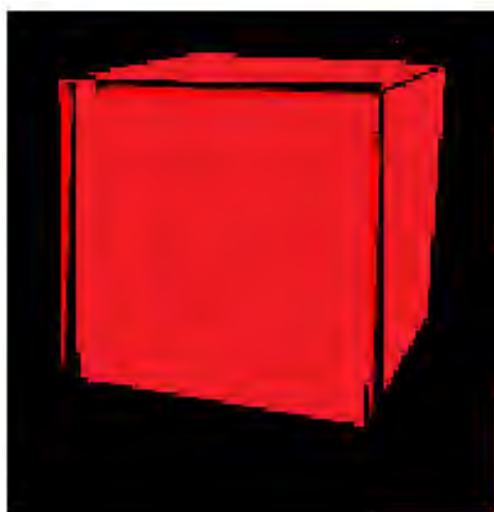


Figure 2.1.1 Initiator for the Sierpinski Gasket (Cube-based) having depth=3.

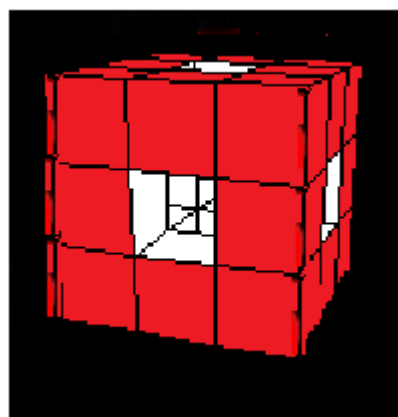


Figure 2.1.2 Generator for the Sierpinski Gasket (Cube-based) having depth=3. The generator is based on the IFS Code as outlined in the end of section 3.

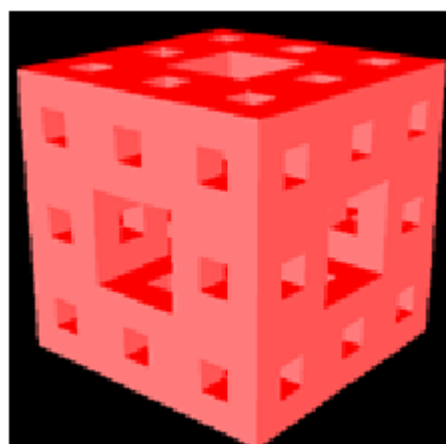


Figure 2.2.1 Original 3D Sierpinski Gasket (Cube-based) having depth=3 obtained from the initiator and generator as shown in Figures 2.1.1 and 2.1.2.

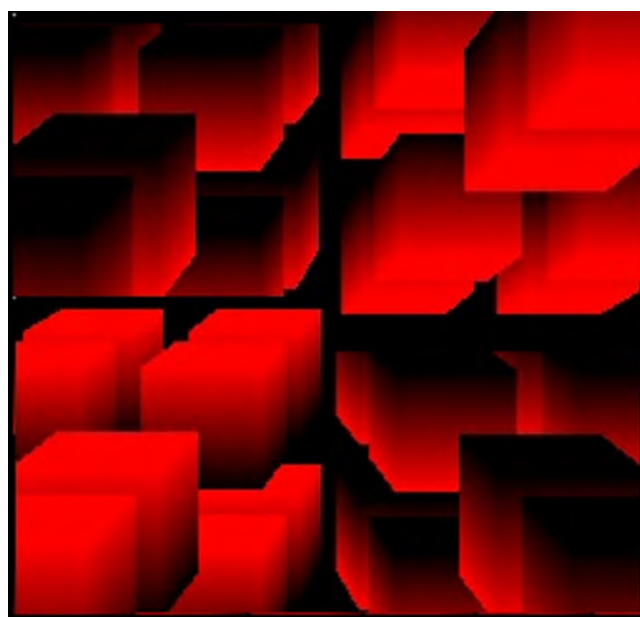


Figure 2.2.2 3D Sierpinski Gasket (Cube-based) having depth=3 obtained by recursive re-generation of Figure 2.2.1 (using the generator as shown in Figure 2.1.2).



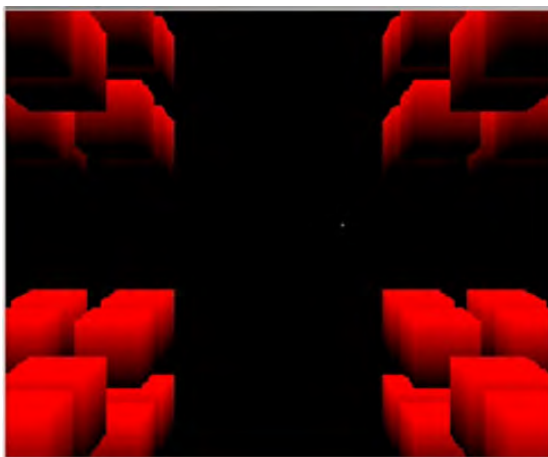


Figure 23 3D Sierpinski Gasket (Cube-based) having depth=3 obtained from recursive re-generation of Figure 2.2.2 (using the generator as shown in Figure 2.1.2).

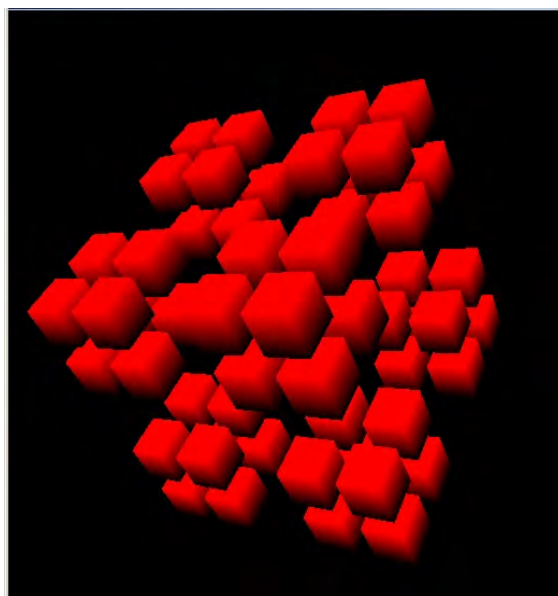


Figure 3 3D Sierpinski Gasket (Cube-based) having depth=3 and applying the generator on Figure 2.3.

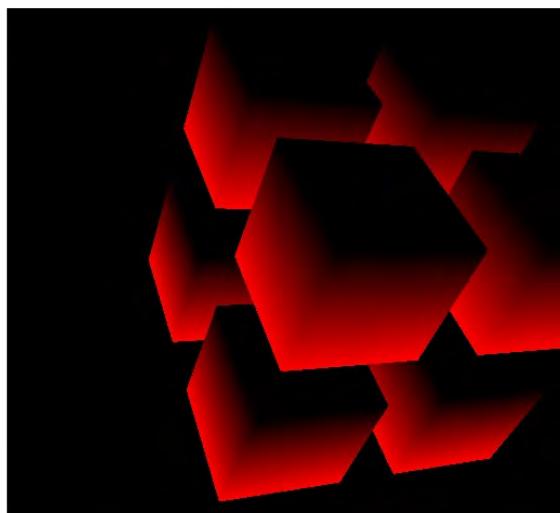


Figure 4 3D Sierpinski Gasket (Cube-based) generated using depth=2 applying the generator on Figure 3.



Figure 5 The 3D Sierpinski Gasket (Cube-based) generated using depth=1 applying the generator on Figure 4.

5. Conclusion

The 3D version of the Sierpinski Carpet termed as Sierpinski Gasket gives an idea about the nature and mathematical aspect of the Sierpinski fractal. This is analogous in perspective to the version of the Sierpinski Triangle known as Sierpinski Pyramid. This paper presents a method of algorithmically maneuvering 3D fractal images that illustrates the concept of closing the self-similarity loop of fractals by way of generating a 3D Sierpinski Gasket fractal starting with a cube as the baseline shape and applying the same algorithm recursively by way of IFS-like transformations of ((x,y) rotation, z (zoom)) – and changing additional variant property called depth (that represents a new property corresponding to projection) from 3 to 2 to 1 in that order. The source image for the 3D Gasket is generated based on the 2D version of the same that is obtained by applying IFS transformations to a square figure after $n \geq 3$ iterations where $n \geq 3$. The displayed output of the same based on a typical set of inputs is presented. The experimental results are shown in section 4 and concluding remarks in section 5. The results demonstrate the closeness in self-similarity aspect of fractal images and how by using theory, computation and experimentation that fractal geometry is both an art and a science.

6. Acknowledgements

Our thanks to all involved for their guidance and valuable suggestions especially to my guide Dr. Jibitesh Mishra for his occasional advice.



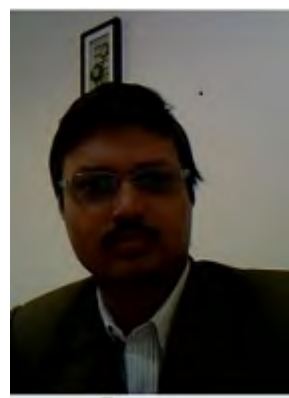
7. References

- [1] Yokoya, N., Yamamoto, K. and Funakubo, N.: Fractal-based analysis and interpolation of 3D natural surface shapes and their application to terrain modeling. *Comp. Vis., Graph. Image Proc.*, 46 (1989) 289–302.
- [2] Hart, J. C. and DeFanti, T. A.: Efficient antialiased rendering of 3-D linear fractals. *Proc.SIGGRAPH'91, Computer Graphics*, 25(4) (1991) 91–100.
- [3] Wyvill, G. B. and McPheeters, C.: Solid texturing of soft objects. *IEEE Comp. Graph.and Appl.*, 7(12) (1987) 20–26.
- [4] Norton, A.: Generation and display of geometric fractals in 3D. *Computer Graphics*, 16(1) (1982) 61–67.
- [5] Groeller, E. and Wegenkiltl, H.: Interactive design of non-linear functions for iterated function systems. *Proc. 4th Intl. Conf. Comp. Graph. and Visual. WSCG'96, Plzen* (1996) 93–102.
- [6] Barnsley, M. F., Elton, J. H. and Hardin, D. P.: Recurrent iterated function systems. *Constructive Approximation*, 5 (1989) 3–31.
- [7] Barnsley, M. F. and Hutchinson J.: New methods in fractal imaging. *Proc. CGIV'06, Sydney* (2006).
- [8] Nikiel, S.: An efficient fractal modeller (modification of the IFS). *Proc. 4th Intl. Conf. on Comp. Graph.Visual., WSCG'96, Plzen* (1996) 294–300.
- [9] VanWijk, J. J. and Saupe D.: Image-based rendering of iterated function systems. *Computers & Graphics*, 28(6) (2004) 937–943.
- [10] Monro, D. M. and Dudbridge, F.: Rendering algorithms for deterministic fractals. *IEEE Comp. Graph. and Applic.*, 272(17) (1995) 32–41.
- [11] Chen, Y. Q. and Bi, G.: 3D IFS fractals as real-time graphics model. *Computers & Graphics*, 21(3) (1997) 367–370.
- [12] Rama, B. and Mishra, J. 2011. Generation of 3D fractal images for Mandelbrot Set. *ACM-International Conference Proceedings Series Proceedings of the 2011 International Conference on Communication, Computing & Security*, doi>[10.1145/1947940.1947990](https://doi.org/10.1145/1947940.1947990), 235-238.
- [13] Rama, B. and Mishra, J. 2010. Generation of 3D Fractal Images for Mandelbrot and Julia Sets. *Special Issue of Intl. Jour. of Computer and Communication Tech.* 1, 2-4, 178-182.
- [14] Mandelbrot, B., 1982, *The Fractal Geometry of Nature*, W.H. Freeman & Co, New York, NY.
- [15] Mirtchovski. A, 2001, Sierpinski's gasket in 3d (using cubes), url><http://mirtchovski.com/usask/opengl/>
- [16] OOoAuthors Group. 2010. Changing Object Attributes, OpenOffice.org 3 Draw Guide, Chapter 04, url>http://wiki.services.openoffice.org/wiki/Documentation/OOo3_User_Guides/Draw_Guide/Changing_Object_Attributes.
- [17] Bourke. P, 1993. Sierpinski Gasket, url><http://local.wasp.uwa.edu.au/~pbourke/fractals/gasket/>.

Biographies



Bulusu Rama is currently Assistant Professor in the Department of Computer Science & Engineering at MLR Institute of Technology, a constituent college of Jawaharlal Nehru Technological University, Hyderabad, India. He has about 12 years of academic experience in mentoring, teaching and guiding students in Computer Science discipline with special reference to Database Systems, Computer Graphics and Data Structures. He holds a Master of Technology in Computer Science, a Master of Science in Mathematics and also a double Honors graduate in Computer Science & Engineering and Mathematics. He has contributed papers to various International and national conferences and also to technical journals. He can be reached at rama_bulusu@yahoo.com.



Dr Jibitesh Mishra is currently Head of the Department of Information Technology, College of Engineering & Technology, a constituent college of Biju Patnaik University of Technology, Orissa. He has more than 16 years of teaching experience in various universities throughout the world. He has published many books of repute. His first book *Design of Information System: A Modern Approach*, was published in India, UK and China in the year 2000. His research interests are Fractal Graphics and Web Engineering. He has published many research papers in international journals & conference proceedings.





Monitoring intra-urban changes with Hidden Markov Models using the spatial relationships

H. Essid^{1,2}, A. Sellami², V. Barra¹, I.R. Farah^{2,3}

¹*Clermont-Université, Université Blaise Pascal, LIMOS, BP 10448, F-63000 CLERMONT-FERRAND, France*
CNRS, UMR 6158, LIMOS, F-63173 AUBIERE, France

²*Research Laboratory in Computer Integrated Documentation and Arabized- Documentiel geniuses and Software-National School of Computer Science*
Campus Universitaire de Manouba, 2010 Manouba, Tunis, Tunisia

³*TELECOM-Bretagne, Département ITI*
Technopôle Brest Iroise CS 83818, 29238 Brest Cedex, France
 [akrem.sellami, vincent.barra, houcine.essid]@isima.fr, riadh.farah@ensi.rnu.tn
<http://www.isima.fr>, <http://www.ensi.rnu.tn>, <http://www.telecom-bretagne.eu>

Abstract

This paper presents a methodology for integrating a new parameter measuring spatial relationships in the hidden Markov models (HMM) in order to detect, interpret and predict changes in urban areas from satellite images. This approach is divided into three phases: the detection of different spatial relationships in the urban area; the training of a hidden Markov model using Baum-Welch learning algorithm, integrating the changing spatial relationships obtained through the Allen's temporal algebra; the interpretation of changes in urban area and the prediction of future changes.

Simulated spatiotemporal changes on synthetic data show the interest of this method for the analysis of spatiotemporal changes of relations between objects. Results allows detection and prediction to be performed from the various time series of images for the observations of spatiotemporal events such as urban expansion. It is therefore reasonable to use this approach to interpret and estimate the movement of the urban area.

Keywords: *Change detection, Spatial relationships, Allen's temporal algebra, Intra-urban analysis, Hidden Markov model.*

1. Problematic and state of the art

Today, time series of satellites images are easily available and allow the evolution of earth surface to be observed. These series constitute a great volume of data and they contain complex information. For example, many spatiotemporal events, such as the harvest, the maturation of culture or the evolution of urban areas, can be observed and are useful for e.g. the management of land use or the understanding of town expansion.

We propose in this work to monitor intra-urban changes using time series of satellites images.

Several authors [3] [7] [21] have proposed a comprehensive review of existing approaches to change detection in remote sensing. Derrode [23] and Carincotte [6] have developed a method to detect changes between two Synthetic Aperture Radar (SAR) images using fuzzy hidden Markov chains. Bouyahia et al. [28] have also used hidden Markov models to unsupervised change detection in bi-date SAR images. Aach et al. [24] developed bayesian algorithms to change detection using Markov random fields. Antonius et al. [4] developed a technique for Markov change detection that can be used to predict future changes based on the rate of change in the past and generating probabilities of changes between classes.

Oscar et al. [25] [26], Van et al. [9] and Weizman et al. [15], also used information such as the texture or, gray-levels for the segmentation and the recognition of urban areas. Lai et al. [27] have used perimeter, texture mean, and area such as parameters for segmentation of urban area. Finally, Nielsen [18] used an approach (Window Independent Context Segmentation (WICS)) for the extraction of different urban area categories from satellite images.

In the spatio-temporal segmentation domain, studies have mainly focused on the modeling of temporal texture. For example, in [2], authors generalize the method of Haralick et al. [22] using the co-occurrence matrix in three dimensions to extract information related to the evolution of a texture.



Spatial relationships constitute structured knowledge, as opposed to digital information such as gray-level, color, or texture that are extracted from images.

Our goal, therefore, is to present a methodology to monitor, interpret, and predict the intra-urban changes, using Hidden Markov Model (HMM) taking into account the spatio-temporal relationships.

Hidden Markov Models (HMM): Hidden Markov Model (HMM) [8] represents a double embedded stochastic process. In an HMM, the observations (V_i) are regarded as symbols emitted by non observable states (S_i), following particular probabilistic functions, whereby the state sequence is a first order Markov Chain.

Let N be the number of states in the model (the individual states are denoted as $S = \{S_1, \dots, S_N\}$, and q_t be the state of S_i at time t , let M be the number of distinct observation symbols per state (the individual symbols are denoted as $V = \{v_1, \dots, v_M\}$). A basic HMM consists of three sets of parameters:

- The symbol emission probabilities b_{jk} - the probability that symbol v_k is emitted by state S_j

$$b_{jk} = P[v_k \text{ at } t | S_j = q_t] \quad 1 \leq j \leq N, 1 \leq k \leq M$$

- The state transition probabilities a_{ij} - the probability of being in state S_j in the subsequent time instant given that the current state is S_i

$$a_{ij} = P[S_j = q_{t+1} | S_i = q_t], \quad 1 \leq i, j \leq N$$

- The prior probability distribution π_i that the system is in a given state S_i at the initial time instant

$$\pi_i = P[S_i = q_1], \quad 1 \leq i \leq N$$

An HMM is typically written as: $\lambda = \{A, B, \pi\}$.

Spatial relations: The spatial relationships constitute the basis of linguistic description of spatial configurations. These spatial relationships are generally classified into different categories [14]: topological, distance and directional relationships. They represent an effective way to describe structured scenes [12].

The distance relationships are based on the notion of distance in a Euclidean space.

The directional relationships represent the positions of objects one with respect to the other, or relative to the frame of the image, the same way that geographic directions are expressed using four cardinal points [1]. Indeed, there are mainly three types of directional relationships:

- *Strict:* North, South, East and West.
- *Mixed:* North, South, East and West, Northeast, Northwest, Southeast and Southwest.
- *Position:* to the right, to the left, below, above.

Finally, the topological relationships don't give information on the cardinal positioning of objects, but rather interaction between two objects. The model the most widely used is the model of "nine intersections" of Egenhofer [16] [17]. In [19], eight relations as formally defined. Figure 1 that capture the topological relations between a pair of objects.

Relation	Mathematical formulation
Disjoint (A, B)	$A^\circ \cap B^\circ = \emptyset \wedge A^\circ \cap \partial B = \emptyset \wedge \partial A \cap B^\circ = \emptyset \wedge \partial A \cap \partial B = \emptyset$
Meet (A, B)	$A^\circ \cap B^\circ = \emptyset \wedge (A^\circ \cap \partial B \neq \emptyset \vee \partial A \cap B^\circ \neq \emptyset \vee \partial A \cap \partial B \neq \emptyset)$
Inside (A, B)	$A^\circ \cap B^\circ \neq \emptyset \wedge A^\circ \cap B^- = \emptyset \wedge A^- \cap B^\circ \neq \emptyset \wedge \partial A \cap \partial B = \emptyset$
Contains(A, B)	$A^\circ \cap B^\circ \neq \emptyset \wedge A^\circ \cap B^- \neq \emptyset \wedge A^- \cap B^\circ = \emptyset \wedge \partial A \cap \partial B = \emptyset$
CoveredBy(A,B)	$A^\circ \cap B^\circ \neq \emptyset \wedge A^\circ \cap B^- = \emptyset \wedge A^- \cap B^\circ \neq \emptyset \wedge \partial A \cap \partial B \neq \emptyset$
Covers (A, B)	$A^\circ \cap B^\circ \neq \emptyset \wedge A^\circ \cap B^- \neq \emptyset \wedge A^- \cap B^\circ = \emptyset \wedge \partial A \cap \partial B \neq \emptyset$
Equal (A, B)	$A^\circ \cap \partial B = \emptyset \wedge A^\circ \cap B^- = \emptyset \wedge \partial A \cap B^\circ = \emptyset \wedge \partial A \cap B^- = \emptyset \wedge A^- \cap B^\circ = \emptyset \wedge A^- \cap \partial B = \emptyset$
Overlap (A, B)	$A^\circ \cap B^\circ \neq \emptyset \wedge A^\circ \cap B^- \neq \emptyset \wedge A^- \cap B^\circ \neq \emptyset$

Figure 1. Formal definitions of eight relations for topological relation between two objects. A is an object, ∂A is the edge of A , A° its interior and A^- its exterior

Temporal relations: Temporal relations are explicit or inferred relations that are used to describe events or ordered states with respect to time. The main models of temporal relations are Allen's interval algebra [10] [11] and Vilain's point algebra [20]. In Allen's theory 13 relations describe all possible combinations between two intervals (Figure 2).

These basic relations can be combined by disjunction, to obtain 2^{13} relations. In Vilain's point Algebra, three basics relations, above (<), identic (=), and follows (>), are considered.

We are interested here to retrieve different areas (buildings, roads,...) present in urban area, the description of these areas is very interesting and especially the positions and relationships between them. Spatial relationships become an asset to detect these relationships and to describe the urban structural. The use of Allen's interval Algebra allows all the possible relations between two intervals of time to be described and can be integrated in the analysis of time series of images. Now we propose our method, which aims to extract the various urban areas and to detect the different spatial relationships between these areas, and interpret changes through time using the hidden Markov model (HMM).



Relation	Mathematical formulation
i_1 before i_2	$\text{End}(i_1) < \text{Begin}(i_2)$
i_1 meets i_2	$\text{End}(i_1) = \text{Begin}(i_2)$
i_1 overlaps i_2	$\text{Begin}(i_1) < \text{Begin}(i_2) < \text{End}(i_1) \wedge \text{End}(i_1) < \text{End}(i_2)$
i_1 starts i_2	$\text{Begin}(i_1) = \text{Begin}(i_2) \wedge \text{End}(i_1) < \text{End}(i_2)$
i_1 finished i_2	$\text{Begin}(i_1) > \text{Begin}(i_2) \wedge \text{End}(i_1) = \text{End}(i_2)$
i_1 during i_2	$\text{Begin}(i_1) > \text{Begin}(i_2) \wedge \text{End}(i_1) < \text{End}(i_2)$
i_1 equal i_2	$\text{Begin}(i_1) = \text{Begin}(i_2) \wedge \text{End}(i_1) = \text{End}(i_2)$
i_1 contain i_2	$\text{Begin}(i_1) < \text{Begin}(i_2) \wedge \text{End}(i_1) > \text{End}(i_2)$
i_1 finished i_2	$\text{Begin}(i_1) < \text{Begin}(i_2) \wedge \text{End}(i_1) = \text{End}(i_2)$
i_1 started i_2	$\text{Begin}(i_1) = \text{Begin}(i_2) \wedge \text{End}(i_1) > \text{End}(i_2)$
i_1 overlapped i_2	$\text{Begin}(i_2) < \text{Begin}(i_1) < \text{End}(i_1) \wedge \text{End}(i_1) > \text{End}(i_2)$
i_1 met i_2	$\text{Begin}(i_1) = \text{End}(i_2)$
i_1 after i_2	$\text{Begin}(i_1) > \text{End}(i_2)$

Figure 2. Allen's temporal relationships [5]. $\text{Begin}(i)$ and $\text{End}(i)$, represent the start time and end time of intervals

2. Proposed approach

The method we propose is composed of three phases. The first phase is a processing step that aims to build feature vectors, that carry out the evolution of spatial relationships through time.

The second phase is the learning step which aims to train the proposed model using the Baum-Welch algorithm in order to detect changes. Finally, the third phase is the recognition step that allows changes to be interpreted and predictions to be made.

Let $\lambda = \{A, B, \pi\}$ be the hidden Markov model, with A is the transition probabilities; B is the emission probabilities and π the initial probabilities.

Figure 3 shows a generic model for such a HMM.

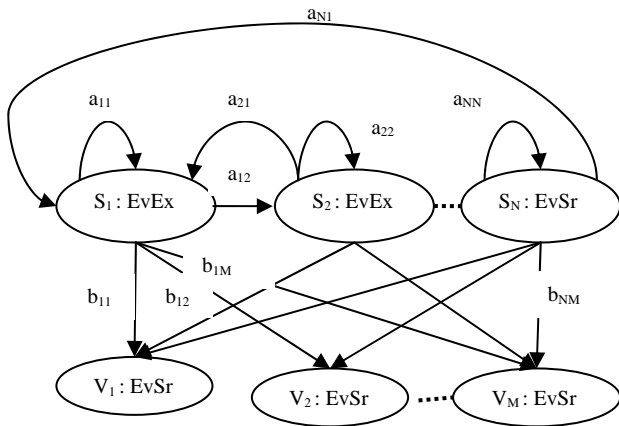


Figure 3. A generic model of hidden Markov model. S is the state set (Evolution of Expansion is denoted as $EvEx$); V is the observations set (Evolution of Spatial relationships is denoted as $EvSr$), a_{ij} is the state transition probability, and b_{jk} is the symbol emission probability

Processing module: Initial data are images acquired at different dates. We segment these images to extract the

different regions of interest, and then detect the different spatial relations between these regions.

Finally, we use Allen's interval algebra to model temporal intervals, and we monitor the evolution of spatial relationships through time. We finally store the information obtained as they are the observations in our hidden Markov model (Figure 4).

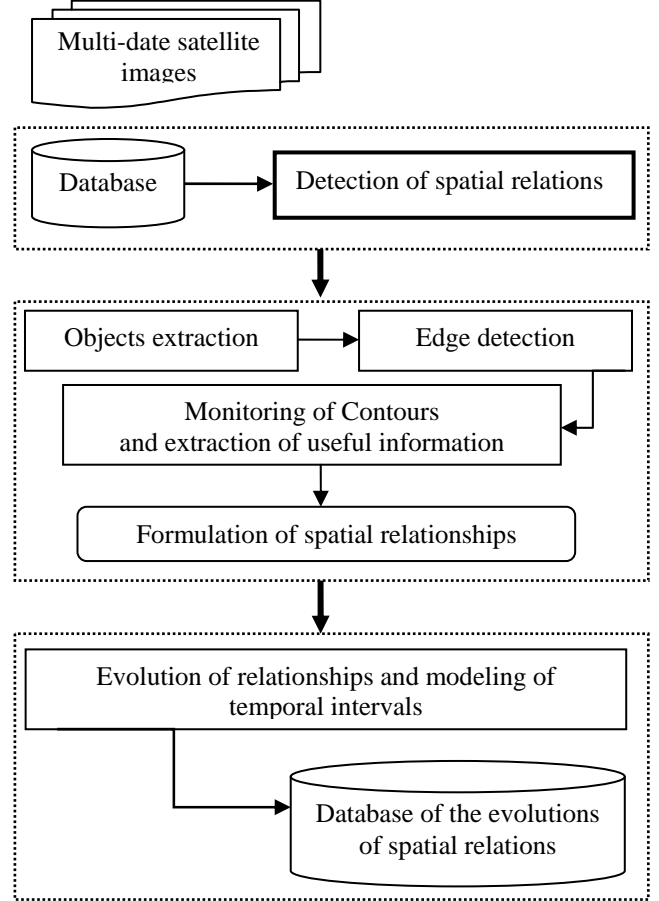


Figure 4. Processing module

Once spatial relationships between regions are detected and monitored, the results of their temporal evolutions are obtained. We then use learning algorithms to train the model using a series of multi-date satellites images.

Learning module: This module aims at estimating the parameter vector λ of the HMM on the basis of a set of sequences of observations called training corpus. Recall that the sequences of observations are the evolutions of spatial relationships over time. The learning algorithm chosen is the Baum-Welch algorithm [13], since it allows a significant improvement compared to traditional learning algorithms. Its goal is to maximize the likelihood of a model λ , it substantially changes the model parameters studied to increase the likelihood. The algorithm performs its optimization by re-estimating the parameters (A , B and π), of the observed sequence.

Figure 5 shows the general process of learning system.



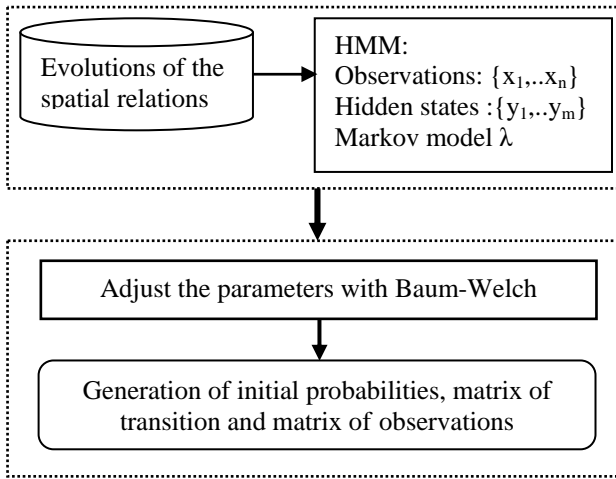


Figure 5. Learning module. $x(i)$ represent the temporal evolutions of spatial relationships and $y(i)$ represent the expansions of regions.

3. Results and discussion

To validate our methodology, we used a series of images which are constituted of two regions. In our case we will identify the different relationships between two binary classes. In a preliminary study, we use synthetic images containing two objects of different colors.

Figure 6 shows some synthetic images containing objects of different colors that change through time.



Figure 6. Synthetic images

Detection of spatial relationships and monitoring their evolutions over time: Table 1 presents the results of experimental trials for some images. For each image at two distinct instants, we present the spatial relationship detected, the temporal evolutions of spatial relationships, and the evolutions of intervals based on the Allen's algebra of intervals.

In order to compute the evolution of spatial relations, we use the index $EvSr$ as an index for computing changes in a spatial relationship between two instants t_i and t_{i+1} .

$EvSr$ = Temporal relation [spatial relationship (t_i), spatial relationship (t_{i+1})]

For example:

If [spatial relationships (t_i) = 'DISJOINT' and spatial relationships (t_{i+1}) = 'OVERLAP']

Then $EvSr$ = 'Strong Evolution'

Instants (i)	Relation (A, B)	EvSr	Modeling of intervals
[1998, 2000]	A disjoints B	Low evolution	i_1 meets i_2
[2000, 2002]	A Meets B		
[2001, 2004]	A Meets B	No evolution	i_2 overlap i_3
[2005, 2006]	A Overlaps B	Strong evolution	i_3 disjoint i_4
[2005, 2010]	A Covers B	Middle evolution	i_4 starts i_5

Table 1. Results of spatio-temporal relations

In this case, we can model the expansion of regions according to the evolution of spatial relationships.

Modeling with HMM:

To use the Hidden Markov Model, observable and hidden states have to be defined. Therefore, we will consider here the temporal evolutions of spatial relationships as the observable states and the expansions of regions as the hidden states.

We propose the following HMM:

- Observable states: {'A': no evolution', 'B': low evolution', 'C': middle evolution', 'D': strong evolution'}
- Hidden states: {'1': no expansion', '2': low expansion', '3': middle expansion', '4': strong expansion'}

We introduce the indices obtained which are the observable states in HMM, and the indices will be ranked according to rule:

- If $EvSr$ is null then no expansion
- If $EvSr$ is low then low expansion
- If $EvSr$ is middle then middle expansion
- If $EvSr$ is strong then strong expansion

Learning phase

The results of the evolution of spatial relationships will be considered as the observations of the HMM. From these observations, we must determine the hidden states associated with them depending on the evolution of spatial relationships. The observable states and the hidden states will be used as the learning parameters of the model. This learning is achieved through the use of the Baum-Welch algorithm.

To train our model, we used the alphabet {'A', 'B', 'C', 'D'}, which represent the evolution of spatial relationships that are observable states in the hidden Markov model, so we used a series of 102 synthetic images, where each image contains two objects of different colors, and these objects change through time. We detected the different spatial relationships, and we got their evolutions through time, and finally we built the observations from the basis of changes in spatial relationships. We obtained the following observations sequence for this series of synthetic images and we introduced them in the learning phase for our system:

"BAAABAABABABBBDBABBAABAAAABDDDBBA
BBABCBAAAABBAABCCBAAAAAABBBBABBAA
AAAABABDCBADAABBBADCABBCBBBDBACAA
B"

After the integration of this sequence in our model we obtained the following results:

The initial probabilities:

1	2	3	4
0,152	0,721	0,027	0,000

The transition matrix:

	1	2	3	4
1	0,2861	0,7139	0,0000	0,0000
2	0,0000	0,1150	0,2750	0,6100
3	0,5509	0,3378	0,0000	0,1123
4	0,0000	0,0046	0,3712	0,6242



The emission matrix:

	1	2	3	4
A	0,3129	0,0000	0,0000	0,9694
B	0,0000	1,0000	0,8672	0,0016
C	0,4337	0,0000	0,0000	0,0290
D	0,2534	0,0000	0,1328	0,0000

• Interpretation phase

Our methodology is now able to interpret the spatio-temporal events of regions from the HMM. Recognition consists in finding the most likely sequence of hidden states leading to the generation of a given output sequence. This sequence of hidden states is generated using the Viterbi algorithm [13].

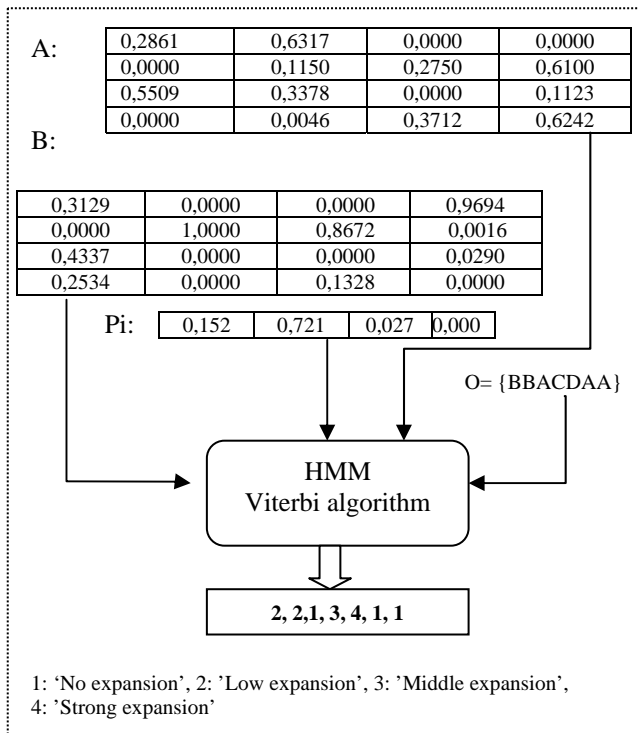


Figure 7. Interpretation of spatio-temporal event

Figure 7 shows an example of spatio-temporal analysis for multi-date images, the recognition of different events (no expansion, low expansion, middle expansion, strong expansion) leading to the generation of the output sequence. We introduced the following sequence of observable states "BBACDAA" which represents the evolution of the spatial relationships of 8 synthetic images in order to interpret changes this series of synthetic images.

Using the Viterbi algorithm we were able to interpret changes in this series of images, where the sequence of hidden states is the expansion of area, so we obtained this sequence of hidden states "2213411", thus, each number in this sequence is the expansion of regions at time t.

• Prediction Phase

To predict the most probable event at a future time, we can use the evaluation function included in the interpretation system. For example, if we want to predict the most probable event at $t = 9$, considering the sequence of observations "BBCDADDB" we first have to compute the probability of observing sequences "BBCDADDBA", "BBCDADDBB", "BBCDADDBC", "BBCDADDBD", then to distinguish the series of the most probable observation and therefore we can predict the event time $t = 9$ using the recognition step of our interpretation system.

Observation O_9	sequence	Logarithm of probability	Probability of E_9	
A	BBCDADDBA	-50.212	strong	0.232
			middle	0.412
			low	0.061
			null	0.295
B	BBCDADDBB	-58.453	strong	0.129
			middle	0.850
			low	0.021
			null	0.000
C	BBCDADDBC	-52.761	strong	0.000
			middle	0.431
			low	0.569
			null	0.000
D	BBCDADDBD	-57.341	strong	0.000
			middle	0.693
			low	0.307
			null	0.000

Table 2. Example of results of events prediction

Table 2 shows an example of events prediction, we notice that this step allows us to predict the states of area expansions. The class B is the most probable and we can predict the state of this region.

4. Conclusion

The proposed methodology allows the spatiotemporal changes in an urban area to be detected from a time series of satellites images. That allows the modeling of changes in urban area and the further interpretation in order to predict future changes. We initially detect spatial relationships to describe structurally the urban area. We modeled the spatio-temporal relations through time with Allen's temporal algebra. The variations obtained are stored in a database. By introducing these changes we were able to interpret changes urban area, and to predict future changes.

Several research axes emerge from this work. We will try in future works to improve the performance of this methodology and to extend our approach of spatio-temporal analysis by adding other indicators of change, since in our methodology we used a single source of information that is spatial relationships. Other indicators can also be used, as factors of change, especially demographics data, energy consumption (electricity, gas, water,...). In this case, we can use others models such as the Coupled Hidden Markov Model (CHMM), since it allows interpretation and the prediction to be performed on the basis of two sources of information that are processed simultaneously.



5. References

- [1] A. U. Frank. Qualitative spatial reasoning about distances and directions in geographic space. *Journal of Visual Languages & Computing*, 3,343-371, 1992.
- [2] A. Rahman and M. Murshed. Real-time temporal texture characterization using block-based motion co-occurrence statistics, *IEEE International Conference on Image Processing*, Vol. 3, Singapore, pp. 1593–1596, 2004.
- [3] A. SINGH. Digital change detection techniques using remotely-sensed data, *International Journal of Remote Sensing* 10, pp. 989-1003, 1989.
- [4] B. Antonius and A. Wijanarto. Application of Markov change detection technique for detecting Landsat ETM derived land cover change over Banten Bay, *Jurnal Ilmiah Geomatika* Vol. 12, No. 1, pp. 11-21, Augustus 2006.
- [5] C. Claramunt and B. Jiang. A representation of relationships in temporal spaces. IN Atkinson , P.& Martin, D. (Eds.) *Innovations in GIS VII: GeoComputation*. London, Taylor & Francis, 2000.
- [6] C. Carincotte, S. Derrode and S. Bourennane. Unsupervised change detection on SAR images using fuzzy hidden Markov chains, *IEEE Trans. on Geoscience and Remote Sensing*, Vol. 44(2), pp. 432-441, February 2006 .
- [7] D. Lu, P. Mausel, E. Brondizio, E. and E. Moran. Change detection techniques, *International Journal of Remote Sensing*, Vol. 25, N°. 12, 2365–2407, 20 June, 2004.
- [8] H. Bunkle and T. Caelli. *Hidden Markov Models – applications in computer vision*, World Scientific, 2001.
- [9] I. Van de Woestyne, M. Jordan, T. Moons and M. Cord. A software system for efficient DEM segmentation and DTM estimation in complex urban areas. In *IAPRS Vol. XXXV, Part. B3, XX ISPRS Congress*, Istanbul, Turkey, 2004
- [10] J, F. Allen. An interval-based representation of temporal knowledge. In *Proceedings of the 7th International Joint Conference on Artificial Intelligence (IJCAI'81)*, pp 221-226, 1981.
- [11] J, F. Allen Maintaining Knowledge about Temporal Inetrvals. *Communications of the ACM*, 26(11):832-843, 1983.
- [12] J,M. Keller and X, Wang. A fuzzy rule-based approach to scene description involving spatial relationships. *Computer Vision and Image Understanding*, 80 :21–41, 2000.
- [13] L. Rabiner. A tutorial on hidden Markov models and selected applications in speech recognition, *Proc. IEEE* 77 (2) pp. 257–286, 1989.
- [14] L. Vieu. Spatial representation and reasoning in artificial intelligence. In O. Stock, éditeur, *Spatial and Temporal Reasoning*, pp. 5-41, Dordrecht, Kluwer, 1997.
- [15] L. Weizman, L and J. Goldberger. Urban-Area Segmentation Using Visual Word, *Geoscience and Remote Sensing Letters*, IEEE, Vol. 6, Nu. 3, pp. 388-392, July 2009.
- [16] M, J. Egenhofer and D,M. Sharma. A critical comparison of the 4-intersection and 9-intersection models for spatial relations: Formal analysis, *Journal Geoinformatica*, Vol 4(4), 1993.
- [17] M,J. Egenhofer, M. Davidand J;H. Mark. The 9-intersection : Formalism and its uses for natural-language spatial predicates.. *Tech. Rep.* 94-1. 1994.
- [18] M. Nielsen.. Extraction of different urban area categories from satellite images using Window Independent Context Segmentation, *Urban Remote Sensing Event (JURSE)*, 2011 Joint , pp 101-104, 2011.
- [19] M. Schneider and T. Behr. Topological relationships between complex spatial objects. *ACM Transactions on Database Systems*, 31(1): 39-81, 2006.
- [20] M. Vilain. and H. Kautz. Constraint propagation algorithms for temporal reasoning. In *AAAI-86 Proceedings*, pp 377-382, 1986.
- [21] P. Coppin, I. Jonckheere, K. Nackaerts. and B. Muys. Digital change detection methods in ecosystem monitoring: a review. *International Journal of Remote Sensing*, 25, pp. 1565-1596, 2004.
- [22] R. Haralick, K. Shanmugan, and I. Dinstein. Textural Features for Image Classification, *IEEE Transactions on Systems, Man and Cybernetics* 6(3) : 610–621, 1973.
- [23] S. Derrode and C. Carincotte. Applications des chaînes de Markov floues à la segmentation non supervisée d'images, *Ateliers Traitement et Analyse de l'Information: Méthodes et Applications TAIMA'05*, pp. 193-198, Hammamet (Tunisia), 26 sept.-1 oct. 2005.
- [24] T. Aach, and A. Kaup. Bayesian algorithms for change detection in image sequences using Markov random fields. *Sig. Proc.: Im. Comm* 7(2):147–160, 1995.
- [25] V, C. Oscar. PhD thesis *Analyse du milieu urbain par une approche de fusion de données satellitaires optiques et radar*. Nice-Sophia Antipolis University, 2003.
- [26] V, C. Oscar, D. Xavier and J. Zerubia. Analyse intra-urbaine à partir d'image satellitaires par une approche de fusion de données sur la ville de Mexico, search report, *INRIA N° 4578*, 2002.
- [27] X. Lai., Y. Wan and W. Wei. A Flow to Generate DEM and Segment Building in Urban Areas from LiDAR Data. In *Proceedings of the XXIst ISPRS Congress*, Beijing, China, 3–11 ; Vol (XXXVII), Part B3b, pp. 149-152, July 2008.
- [28] Z. Bouyahia, L. Benyoussef and S. Derrode. Change detection in synthetic aperture radar images with a sliding hidden Markov chain model, *J. of Applied Remote Sensing*, Vol. 2(1), 023526, doi:10.1117/1.2957968, 2008.



Biographies



Houcine ESSID is a PhD student in computer science from E.N.S.I. Tunisia. Member of RIADI laboratory in Tunisia and LIMOS laboratory in France. houcine.essid@isima.fr



akrem.sellami@isima.fr

Akrem Sellami is a Master student in computer science from FSJEGJ, University of Jendouba, Tunisia. He is a member at laboratory RIADI, ENSI National School of Computer Sciences engineering, Tunisia, (since 2011).



Vincent Barra is full Professor in Computer Sciences at the Blaise Pascal University, Clermont-Ferrand, France. His research interests include image processing, statistical learning and 3D data analysis. He's author or co-author of more than 20 scientific papers and 80 papers in proceedings of national or international conferences. He participates as a PC member to several national and international conferences. vincent.barra@isima.fr

Vincent Barra is full Professor in Computer Sciences at the Blaise Pascal University, Clermont-Ferrand, France. His research interests include image processing, statistical learning and 3D data analysis. He's author or co-author of more than 20 scientific



Imed Riadh Farah received the M.D. and Dr Eng. degrees, from ISG Inst. of computer sciences in 1995 and ENSI Sch, in 2003, respectively. After working as a research assistant (from 1996) and a permanent researcher at laboratory RIADI, ENSI National School of Computer Sciences engineering (since 1995), he has been an assistant professor at University of Jendouba, since 2004. His research interest includes image processing, pattern recognition, artificial intelligence, data mining and their application to remote sensing. He is a member of Arts-Pi Tunisia. riadh.farah@ensi.rnu.tn

Imed Riadh Farah received the M.D. and Dr Eng. degrees, from ISG Inst. of computer sciences in 1995 and ENSI Sch, in 2003, respectively. After working as a research assistant (from 1996) and a permanent researcher at

